Eyes-Free, Touch-Free, and Worry-Free: Repurposing Wireless Chargers for Safer In-Car Gesture Interaction

WEIYI WANG, Shanghai Jiao Tong University, China
LANQING YANG[†,*], Shanghai Jiao Tong University, China
GUO YU, Shanghai Jiao Tong University, China
QINFEI FAN, Shanghai Jiao Tong University, China
JEFFERY WU, Shanghai Jiao Tong University, China
LONGYUAN GE, Shanghai Jiao Tong University, China
LEO LINQIAN GAN, Shanghai Jiao Tong University, China
GUANGTAO XUE[†], Shanghai Jiao Tong University, China

As vehicles evolve into digital living spaces, the demand for seamless human-vehicle interaction has surged. However, traditional interaction modalities impose critical safety and usability burdens. Touchscreens require visual attention (eyes-on), compromising driving safety, while voice assistants, although eyes-free, suffer from inherent latency, high cognitive load, and the intrusive disruption of cabin audio and social conversations. Mid-air gestures offer a promising shortcut for silent, immediate, and eyes-free control. Yet, enabling robust gesture sensing remains a challenge. Vision-based solutions can raise privacy and illumination concerns in some in-cabin deployments; millimeter-wave radar often increases hardware cost and integration complexity; and acoustic sensing can be affected by cabin noise and multipath interference in the in-cabin environment. To bridge this gap, we present *MagHarp* , a novel system that repurposes the ubiquitous in-car wireless charging pad for active gesture recognition. We observe that the human hand acts as a capacitive bridge within the near-field region, subtly modulating the charging signals. By treating the existing Wireless Power Transfer (WPT) hardware as a near-field electromagnetic sensor, we avoid modifying the charger or smartphone. *MagHarp* captures these minute impedance perturbations without modifying the transmitter. We design a robust pipeline featuring a lightweight CNN-BiLSTM network that helps distinguish driver gestures from vehicle-related disturbances and environmental interference. Extensive experiments across 5 chargers and 4 phones demonstrate that *MagHarp* achieves an average gesture recognition accuracy of 97.6% on a 3,500-sample dataset, effective up to 20 cm (recommended within 15 cm for robustness).

CCS Concepts: • **Human-centered computing → Gestural input**.

Additional Key Words and Phrases: In-car Human–vehicle Interaction, Mid-air Gesture Recognition, Repurposed Wireless Charger, Near-field Electromagnetic Sensing

[†]Co-author equal contribution.
[*]Corresponding author.

Authors' Contact Information: Weiyi Wang, Shanghai Jiao Tong University, Shanghai, Shanghai, China, weiyi_wang@sjtu.edu.cn; Lanqing Yang (corresponding author), Shanghai Jiao Tong University, Shanghai, Shanghai, China, yanglanqing@sjtu.edu.cn; Guo Yu, Shanghai Jiao Tong University, Shanghai, Shanghai, China, hugoyu2024@sjtu.edu.cn; Qinfei Fan, Shanghai Jiao Tong University, Shanghai, Shanghai, China, fanqinfei@sjtu.edu.cn; Jeffery Wu, Shanghai Jiao Tong University, Shanghai, Shanghai, China, jeffery2019@sjtu.edu.cn; Longyuan Ge, Shanghai Jiao Tong University, Shanghai, Shanghai, China, gly2000@sjtu.edu.cn; Leo Linqian Gan, Shanghai Jiao Tong University, Shanghai, Shanghai, China, leo-gan@sjtu.edu.cn; Guangtao Xue, Shanghai Jiao Tong University, Shanghai, Shanghai, China, gt_xue@sjtu.edu.cn.



This work is licensed under a Creative Commons Attribution 4.0 International License.
© 2026 Copyright held by the owner/author(s).
ACM 2474-9567/2026/9-ART164
https://doi.org/10.1145/3831984

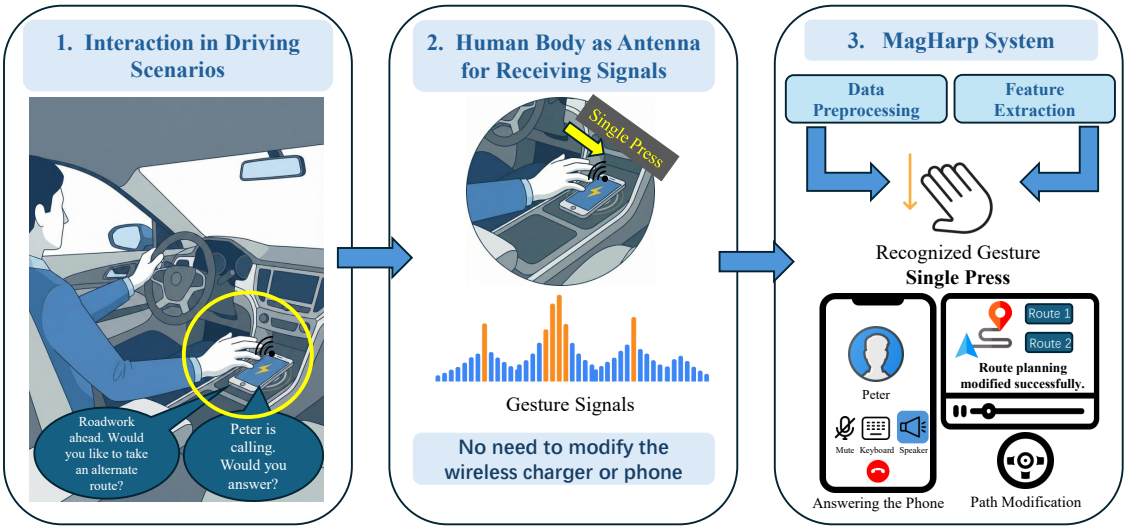


Fig. 1. Schematic diagram of the *MagHarp* system. (Left) During driving, an incoming call notification or navigation route prompt requires attention. (Center) When the driver decides to answer the call or modify the route, they perform a single press gesture in the air, which the *MagHarp* system detects. (Right) After data preprocessing and feature extraction, *MagHarp* classifies the gesture and triggers the corresponding action (e.g., answering the call or modifying the route) on the phone or in-vehicle infotainment system.

ACM Reference Format:
Weiyi Wang, Lanqing Yang, Guo Yu, Qinfei Fan, Jeffery Wu, Longyuan Ge, Leo Linqian Gan, and Guangtao Xue. 2026. Eyes-Free, Touch-Free, and Worry-Free: Repurposing Wireless Chargers for Safer In-Car Gesture Interaction. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 10, 3, Article 164 (September 2026), 34 pages. https://doi.org/10.1145/3831984

1 INTRODUCTION

The modern automobile is undergoing a fundamental metamorphosis, evolving from a simple mode of transport into a "third living space" — a sophisticated, highly connected mobile computing environment. Drivers today are no longer just operating a steering wheel; they are simultaneously managing a complex digital ecosystem. This ecosystem spans two distinct but overlapping worlds: the Vehicle Infotainment System (handling navigation [2], climate, and media [15]) and the Personal Mobile Device (handling communication [41], social updates, and on-demand service workflows [3, 16, 34, 36]). Whether it is a commuter skipping a song on the car stereo, a ride-hailing driver confirming an order within a 15-second window [48], or a user dismissing a distracting notification, the demand for interaction is continuous, high-frequency, and cognitively demanding [7].

This explosion of functionality has created a critical "Interaction-Safety Paradox". While the digital throughput in the cabin increases, the driver's cognitive and visual bandwidth remains strictly limited. To bridge this gap, the industry has introduced various interaction modalities, yet each involves practical trade-offs in the driving context. Touchscreens, currently the dominant interface, demand precise hand-eye coordination, inevitably forcing eyes-off-road moments that compromise safety [1, 37, 40]. Voice Assistants, while theoretically eyes-free, suffer from high latency, struggle in noisy cabin environments, and introduce social friction — passengers often find it awkward or disruptive when drivers shout commands to mute a call or change a route [27]. To address these trade-offs, Contactless Gesture Interaction has emerged as a promising complementary interface, offering eyes-free, immediate command execution. However, deploying gesture sensing robustly in real vehicles remains nontrivial. Vision-based systems (RGB/IR cameras) provide rich spatial information and have shown strong

performance for in-cabin perception, but their deployment may still involve practical trade-offs related to camera placement, occlusion, illumination control, and privacy expectations in some settings [8, 18, 21, 44].

To bypass these limitations and occlusions, researchers have explored other sensing modalities. RF-based approaches (mmWave, UWB, WiFi) enable privacy-friendly sensing [24, 50–52]. State-of-the-art systems achieve sophisticated tracking [30] and handle complex environments through advanced techniques [5, 19]. Acoustic and Vibration sensing also repurpose existing hardware, with ultrasonic techniques offering fine-grained tracking [31] and vibration signals used for device interaction [13, 14]. These modalities are promising, but they may still involve trade-offs in vehicles: Radar and UWB systems often require dedicated hardware, additional integration effort, and regulatory consideration [47]; RF sensing must carefully handle multipath in metallic cabins [50]; and acoustic methods can be affected by cabin noise and comparatively weaker spatial confinement [31, 45, 46, 57]. In parallel, Electromagnetic (EM) field sensing has shown promise, offering a middle ground with better spatial containment. However, existing EM approaches only partially match the target in-cabin interaction scenario: passive methods depend on leakage sources whose characteristics vary with operating state [12], while active magnetic methods typically rely on wearables or dedicated tags [6, 22, 49]. This leaves a critical gap for an interaction modality that is both infrastructure-based, wearable-free, and robust against environmental interference, ideally located in a natural, ergonomic position within the driver's reach.

We identify this opportunity in the proliferation of Wireless Power Transfer (WPT) Charging Pads. With adoption rates exceeding 50% in new vehicles and over 80% of automotive modules obtaining Qi 2.0 certification [11, 42], the charging pad has established itself as the "Digital Hearth" of the cockpit. It is the default resting spot for the smartphone — physically located in the "Golden Zone" of driver interaction where the hand naturally hovers to access the phone or gear shift [20]. Currently, however, this hardware is utilized as a blind utility, serving solely as a unidirectional power conduit. In traditional power engineering, human interference in the magnetic field is viewed solely as parasitic noise to be suppressed [23].

In this paper, we propose a paradigm shift: we reinterpret this human-induced interference as a rich, exploitable information channel. We introduce *MagHarp* , a system that transforms the existing in-vehicle wireless charging infrastructure into a unified gesture interface as Fig. 1 shows. By analyzing the electromagnetic leakage fields inherently emitted by WPT devices [28, 38], *MagHarp* detects and classifies micro-gestures performed in the proximity of the pad, effectively turning it into an in-air gesture sensing surface. This approach unlocks a dual-interaction capability: the system can control vehicle-native functions (e.g., accepting a navigation reroute) or act as a proxy for device-centric control (e.g., answering a call), all performed at the natural resting location of the driver's hand without physical touch. Our core insight is that the WPT field, while primarily near-field, is highly sensitive to the capacitive coupling of the human body. Unlike RF which suffers from multi-path, or acoustics which are susceptible to ambient noise, the WPT leakage field offers a more contained and stable interaction zone. Furthermore, hardware security research has already demonstrated the physical responsiveness of WPT fields to impedance changes [29, 38, 55], confirming our hypothesis.

Realizing this vision, however, presents three distinct technical challenges: (1) Hostile Signal Characteristics: Unlike dedicated sensing signals, WPT emissions are designed for energy transfer. The signal is dominated by strong carrier waves and switching harmonics, burying the minute modulation caused by hand gestures. (2) Dynamic Operating States: To maintain efficiency, modern chargers continuously adjust their operating frequency and duty cycle in response to battery levels and phone alignment. This creates a non-stationary spectral baseline that confuses traditional motion detection algorithms. (3) Complex Electromagnetic Environment: The vehicle cabin is electromagnetically noisy, making it difficult to isolate human-induced perturbations from environmental clutter.

To overcome these challenges, we introduce a physics-informed impedance decoding framework. We analytically model the human body as a conductive antenna that modulates the coupling impedance of the WPT

environment. Guided by this physical ground truth, we design a signal processing pipeline that isolates gesture-specific harmonics from charging noise and employ a lightweight CNN-BiLSTM network to decode temporal patterns. *MagHarp* is implemented as a passive add-on sensing prototype using a commodity SDR and a low-cost conductive-fabric pickup, while leaving the charger and smartphone unchanged. We evaluate *MagHarp* through extensive experiments involving 10 users, 4 smartphones, and 5 commercial wireless chargers, collecting a total of 3,500 gesture samples. The results demonstrate that our system maintains robust performance with 97.6% accuracy across a 10-day span.

We summarize our major contributions below.

- We present the first contactless interaction scheme that repurposes the electromagnetic leakage from in-vehicle wireless chargers. We envision this "infrastructure-opportunistic" paradigm transforms passive power conduits into active sensing surfaces, enabling unified control of both vehicle systems and smartphones without adding a new active sensing transmitter.
- Through rigorous theoretical analysis, we model the relationship between human hand maneuvers and the modulation of coupling impedance in the WPT near-field. We believe this physics-informed impedance decoding framework provides a fundamental understanding of human-WPT interaction, laying the theoretical foundation for robust feature extraction against dynamic charging states.
- We design a passive add-on sensing architecture that requires no hardware changes to commercial chargers or phones. By combining harmonic-band stitching with a lightweight CNN-BiLSTM network, we effectively suppress environmental interference and charging noise, making the proposed sensing modality practical and cost-effective for mass deployment.
- We conduct comprehensive evaluations across 5 wireless chargers, 4 smartphones, and 10 users involving over 3,500 samples to assess performance under real-world variables. The results demonstrate the system's robustness across diverse devices and cross-day scenarios, validating its suitability for future smart cockpits.

The rest of the paper is organized as follows. We first introduce related work in §2. Then we introduce the background of WPT leakage signals and the human-antenna effect in §3 and analyze its feasibility in §4. Afterwards, we elaborate on the system design of *MagHarp* in §5, including the detailed signal processing and gesture classification pipeline. The implementation details and evaluations are presented in §6 and §7, respectively, followed by discussions (§8) and conclusions (§9).

2 RELATED WORK

2.1 Vision-Based In-Cabin Gesture and Activity Recognition

Vision remains the dominant modality for in-cabin interaction due to its high spatial resolution. Prior research has extensively deployed RGB and IR cameras for tasks ranging from dynamic hand gesture recognition [35] to holistic driver activity understanding [33]. To enhance robustness, recent efforts focus on fine-grained benchmarks like IVGaze for gaze estimation [8], and temporally-aware models that exploit inter-frame attention to capture subtle cues [21]. Large-scale multi-modal datasets like AIDE further facilitate multi-task learning for comprehensive cockpit perception [53], while neuromorphic event cameras offer a potentially more privacy-preserving alternative for motion capture [18, 44].

Limitations: Nevertheless, their deployment in production vehicles can still depend on line-of-sight, mounting geometry, and illumination management, and may raise privacy or data-governance concerns in some settings. These practical trade-offs motivate complementary non-visual interaction channels.

2.2 RF, Acoustic, and Vibration-Based Contactless Interaction

To bypass visual limitations, researchers have explored signals that do not require line-of-sight. RF-based approaches (mmWave, UWB, WiFi) enable privacy-friendly sensing [24, 50, 52]. State-of-the-art systems like

Table 1. Comparison with State-of-the-Art In-Cabin Sensing. Our *MagHarp* enables robust, privacy-safe interaction by reusing existing infrastructure. (✓: Good, ✗: Poor, ~: Limited).

Sensing Modality & Representative Works	Privacy Safe	Light Immunity	Noise Immunity	Spatial Select.[†]	Wearable-Free[‡]
Vision-Based Approaches *RGB/IR:* Gesture [35], Activity [33], Gaze [8, 21]; *Event Cam:* [18, 44]	✗	✗	✓	✓	✓
RF Sensing (Radar / WiFi) *Radar:* Pose [30], Fatigue [5, 58]; *WiFi/UWB:* [24, 56]	✓	✓	✓	✗	✓
Acoustic & Vibration *Audio:* VeCare [57]; *Ultrasound:* [31, 46]; *IMU:* VibWriter [13]	~	✓	✗	✗	✓
Magnetic & Electromagnetic (EM) *Active:* MagTrack [22], MagX [6]; *Security:* MagPrint [54], MagSpy [17]; *Passive:* EVLeSen [12]	✓	✓	✓	~	~
***MagHarp* (Ours)** *Infrastructure-Opportunistic WPT Sensing*	✓	✓	✓	✓	✓

[†] **Spatial Select.**: Physically confines sensing to driver area (vs. broadcast RF/Audio).

[‡] **Wearable-Free**: No body-worn or handheld device is required during interaction.

M3Track achieve 3D posture tracking [30], while SeRadar [19] and MSense [5] employ advanced beamforming to handle complex environments. Acoustic and Vibration sensing repurpose existing hardware for interaction. Ultrasonic techniques enable fine-grained tracking [31], and vibration signals have been innovatively leveraged for handwriting recognition and input on mobile devices [13, 14]. Others use distributed acoustic nodes for cabin monitoring [45, 46, 57].

Limitations: Despite their strong capabilities, these modalities also involve practical trade-offs in vehicles. Radar and UWB can provide robust perception, but often require dedicated RF hardware, packaging space, and additional integration effort. In metallic cabins, multipath remains an important design consideration rather than a solved default. Acoustic approaches can reuse commodity transducers, but their performance may still be affected by cabin noise and comparatively weaker spatial confinement. Voice-based control, specifically, suffers from social interruption, as issuing commands can be awkward and disruptive during conversations.

2.3 Magnetic and EM-Field-Based Ubiquitous Interaction

Electromagnetic (EM) sensing offers a compelling middle ground with better spatial containment. The ubiquity of EM signals has been extensively explored for security and user characterization. For instance, MagPrint [54] utilizes EM signals for deep-learning-based user fingerprinting. Similarly, magnetometers have been used to detect eavesdropping (MagDefender [39]) or reveal privacy leakages (MagSpy [17]), demonstrating the high sensitivity of magnetic channels. In the automotive context, EVLeSen repurposes leaked EM signals from EV powertrains for passive sensing [12], while other works leverage wearable magnetics (MagTrack [22], MagX [6]) or capacitive coupling [49] for robust tracking.

Limitations: While demonstrating physical feasibility, existing EM approaches face a dilemma for in-cabin interaction. Passive methods (like EVLeSen) rely on uncontrolled leakage sources that vary unpredictably with vehicle states. Active magnetic methods (like MagTrack) typically mandate the user to wear extra sensors, failing the wearable-free requirement. This highlights a gap for an interaction modality that is both infrastructure-based and robust against environmental interference.

2.4 Repurposing Wireless Power Transfer (WPT) for Sensing

Wireless Power Transfer (WPT) pads are becoming standard in vehicle center consoles, yet their sensing potential remains under-explored in HCI. However, the physical responsiveness of WPT fields has been validated by hardware security research. Studies like WISERS [38] utilize coil whine and magnetic perturbations to infer smartphone activities, while MagFingerprint [29] confirms the stability of WPT signals for authentication. Recently, VoltSchemer [55] showed that voltage noise propagates through the magnetic field, proving the channel is highly responsive to impedance changes.

Our Approach: We bridge the gap between these physical insights and HCI needs. Unlike security attacks that treat these signals as vulnerabilities, we repurpose the WPT infrastructure as an intentional interaction interface. To ground this approach, we next analyze the physical characteristics of WPT leakage and derive a model for human-induced modulation.

3 PHYSICAL PRINCIPLES AND MODELING

Before detailing our system architecture, we establish the theoretical foundation of *MagHarp* . This section analyzes the electromagnetic characteristics of the Qi protocol, derives an equivalent circuit model for the human-WPT coupling channel, and validates the observability of gesture-induced perturbations.

3.1 Wireless Charging Emissions and Leakage-Signal Modeling

Smartphone wireless charging relies on near-field resonant magnetic coupling: a high-frequency inverter drives the transmitter (Tx) coil to generate alternating magnetic flux, while the receiver (Rx) coil is tuned to a nearby resonance for efficient energy transfer across a millimeter-scale gap. The Tx coil current near the fundamental is:

$$i_{\mathrm{Tx}}(t) = \Re\{I_0(t)e^{j\omega_0 t}\}, \tag{1}$$

where $\omega_0 = 2\pi f_0$ is the nominal angular frequency and $I_0(t)$ varies slowly with power control and load conditions. The Tx–Rx coupling, characterized by mutual inductance M, induces voltage at the Rx following Faraday's law:

$$v_{\mathrm{ind}}(t) = -M\frac{di_{\mathrm{Tx}}(t)}{dt}. \tag{2}$$

Capacitive compensation on both sides achieves approximate resonance around ω_0, canceling reactive components and improving transfer efficiency.

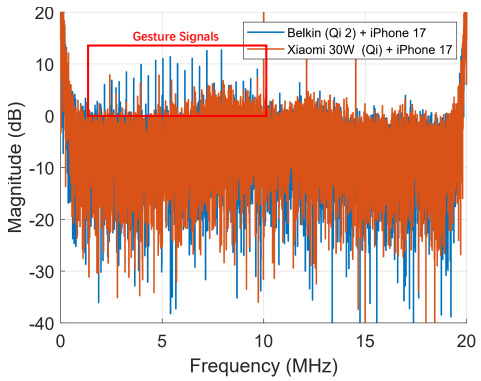
Although power transfer relies on near-field magnetic coupling, practical chargers inevitably generate leakage electromagnetic fields: part originates from the time-varying magnetic dipole near field, while additional components arise from inverter switching, parasitic capacitances, and structural interconnects that introduce high-frequency harmonics. For an observation point at position $\mathbf{r}$, the observable field can be represented as:

$$s(\mathbf{r}, t) = \Re\left\{\sum_{k\in\mathcal{K}} S_k(\mathbf{r}, t)\, e^{jk\omega_0 t}\right\}, \tag{3}$$

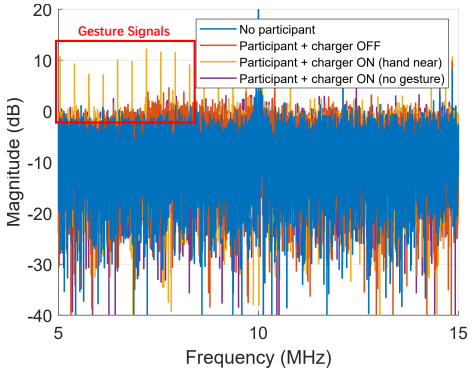
where $k = 1$ corresponds to the fundamental and $k \geq 2$ to harmonics from switching nonlinearities. The *complex envelope* $S_k(\mathbf{r}, t)$ absorbs propagation paths, radiation efficiency, and slow amplitude/phase variations. To capture gesture modulation, we decompose:

$$S_k(\mathbf{r}, t) = \bar{S}_k(\mathbf{r}; \sigma(t)) + \Delta S_k(\mathbf{r}, t), \tag{4}$$

where $\sigma(t)$ denotes the charging state yielding a *slowly varying baseline*, and $\Delta S_k(\mathbf{r}, t)$ denotes the *perturbation* induced by human proximity/motion.



(a) Frequency spectra of Qi and Qi2 wireless chargers measured with an iPhone 17 at 80% state of charge (SoC). Compared with Qi, Qi2 shows higher-amplitude spectral components with lower variability across measurements.



(b) Validation of the human-body-as-antenna effect. Sitting on conductive fabric increases signal amplitude compared to no contact. Hand proximity to the charger creates distinct gesture-related signal patterns.

Fig. 2. Characteristics of Qi charging signals and validation of the human-body-as-antenna effect.

In practice, we select and demodulate a narrowband component around center frequency f_c. Let $k_c \approx \text{round}(f_c/f_0)$ be the closest harmonic index; this component approximates:

$$s_c(\mathbf{r},t) \approx \Re\left\{S_{k_c}(\mathbf{r},t)\, e^{j\omega_c t}\right\}, \quad \omega_c = 2\pi f_c, \tag{5}$$

whose low-frequency envelope variations are carried by $S_{k_c}(\mathbf{r},t)$. The signal received by the human body as a capacitive antenna is essentially this envelope undergoing measurable amplitude/phase perturbations under gesture motion. The following sections model how human–Tx coupling maps into $S_{k_c}(\mathbf{r},t)$, enabling gesture recognition.

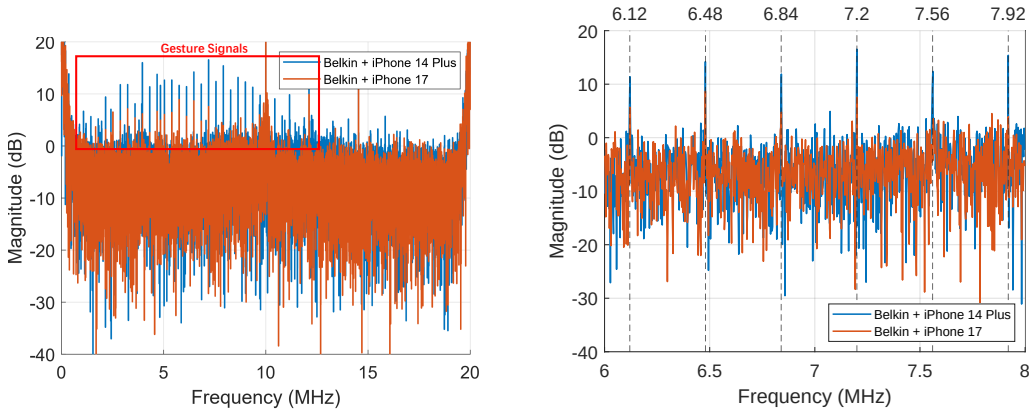
3.2 Wireless Charging Protocols and Signal Characteristics

Legacy Qi (v1.x) operates in the 110–205 kHz band [4], with closed-loop control adjusting inverter frequency and duty cycle in response to receiver feedback [26]. Consequently, dominant spectral components drift over time. By contrast, Qi2 introduces the Magnetic Power Profile (MPP) with magnetic alignment and a fixed operating frequency $f_0 \approx 360$ kHz [10, 26]. This spectral stationarity is advantageous for narrowband filtering, coherent demodulation, and time–frequency feature extraction.

To verify these characteristics, we measured the on-device spectrum using two chargers (legacy Qi and Qi2) with an iPhone 17 at 80% battery level (Fig. 2a). Legacy Qi shows spectral drift across the operating band, whereas Qi2 exhibits a stationary fundamental near 360 kHz with distinct higher-order harmonics at integer multiples. This stability enables consistent bandpass selection and feature alignment across sessions, making Qi2 ideal for gesture recognition.

3.3 Modeling the Human-WPT Coupling Channel

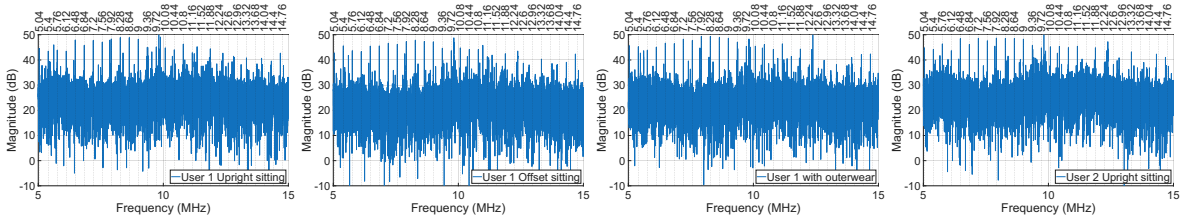
When a human hand approaches a WPT system, parasitic capacitances form between the body and the Tx coil/charger structure, inducing measurable variations in body potential. This "human antenna" effect is well-established in HCI literature [9, 43]. Critically, we measure body potential rather than rectified Rx voltage: participants sit on a conductive-cloth electrode connected via a high-input-impedance amplifier to a HackRF, providing an approximately open-circuit measurement relative to instrument ground.



(a) 0–20 MHz spectrum of two Tx/Rx combinations, showing comb-like harmonics spaced by ~360 kHz; components around ~7 MHz.

(b) Zoom-in of 6–8 MHz. Gesture-related signal modulations are observable on the ~360 kHz harmonics.

Fig. 3. Measured spectral characteristics of Qi charging signals for two Tx/Rx pairs: (a) 0–20 MHz wideband spectrum, (b) zoom-in view of 6–8 MHz highlighting 360 kHz-spaced harmonics used for gesture sensing.



(a) User 1, upright sitting

(b) User 1, offset sitting

(c) User 1, with outerwear

(d) User 2, upright sitting

Fig. 4. Spectral comparisons under representative human-coupling conditions. We compare posture/action variation (upright vs. offset sitting), clothing variation (with outerwear), and user variation (User 1 vs. User 2). These factors slightly affect the baseline amplitude and coupling strength, but the changes are not significant enough to disrupt observability: the 360 kHz-spaced harmonic structure remains present, and gesture-related spectral features are visible across all conditions.

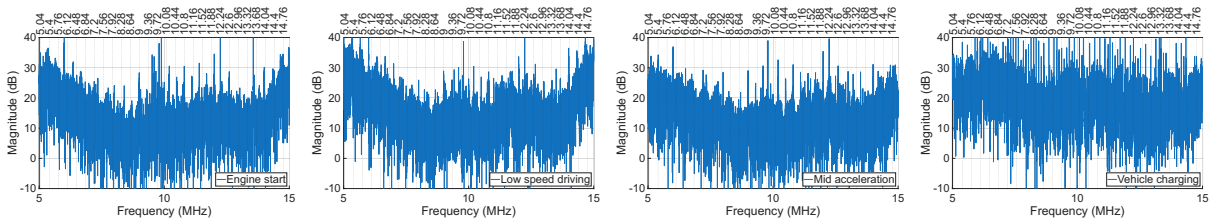
The mapping from leakage field to body/electrode voltage can be abstracted as an equivalent coupling channel varying slowly with posture and environment, capturing parasitic capacitances and input impedance. The narrowband body potential is:

$$v_H(t) \approx \Re\left\{\alpha(t)\, S_{k_c}(\mathbf{r}_H, t)\, e^{j\omega_c t}\right\} + n(t), \tag{6}$$

where $\mathbf{r}_H$ is the effective body location, $\alpha(t)$ is a coupling coefficient varying with posture, distance, and grounding, and $n(t)$ accounts for noise. Equation (6) indicates that the demodulated baseband (complex IQ from HackRF) directly tracks envelope variations of $S_{k_c}(\mathbf{r}_H, t)$. Gesture recognition thus relies on amplitude/phase perturbations induced by body motion.

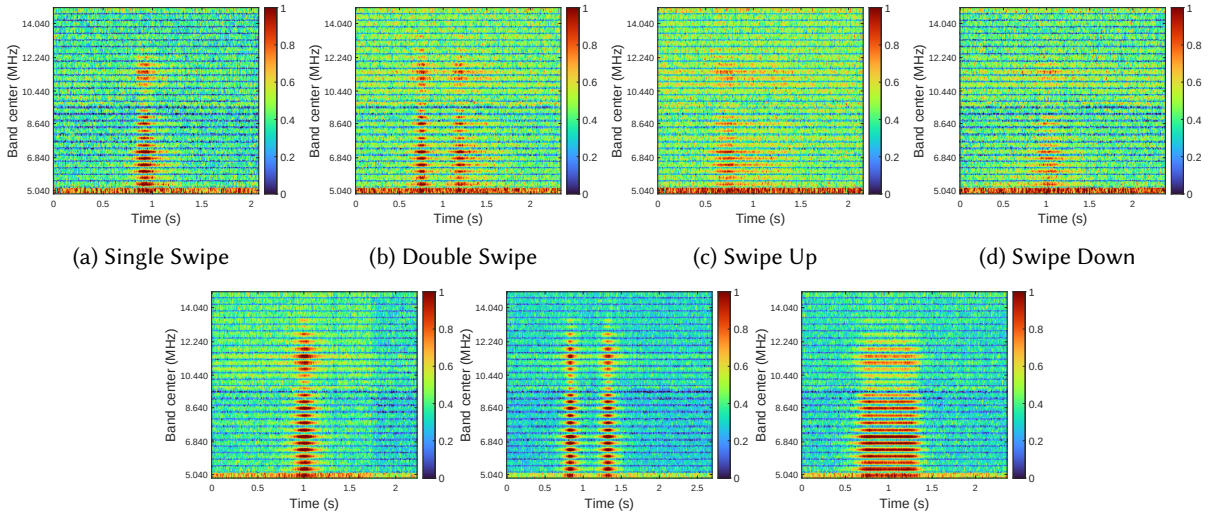
4 FEASIBILITY STUDY

To examine whether WPT leakage can serve as a stable body-coupled sensing channel for in-car gestures, we conduct a feasibility study around four questions. First, we verify whether a user seated on the conductive-fabric



(a) Engine start (b) Low-speed driving (c) Mild acceleration (d) Vehicle charging

Fig. 5. Spectral comparisons under four representative in-vehicle operating conditions: engine start, low-speed driving, mild acceleration, and vehicle charging. Except for vehicle charging, these operating states only slightly perturb the sensing baseline and introduce limited interference to the selected 360 kHz-spaced narrow harmonic bands. Vehicle charging produces stronger noise and contaminates a small number of harmonic bands; nevertheless, several clean and separable harmonic bands remain available for extracting gesture-related spectral features.



(a) Single Swipe (b) Double Swipe (c) Swipe Up (d) Swipe Down

(e) Single Press (f) Double Press (g) Long Press

Fig. 6. Time–frequency visualizations of representative gestures. Different gestures exhibit distinct energy distributions and spectro-temporal shapes, indicating separable patterns.

electrode can receive wireless-charging leakage through the human-body coupling path, and whether bringing the hand close to the wireless charger produces measurable spectral perturbations. Second, we identify stable and high-SNR harmonic components that can be used as opportunistic sensing carriers. Third, we study whether common human-factor variations, including different sitting postures/actions, clothing, and users, significantly change the sensing baseline. Finally, we evaluate whether representative in-vehicle operating states introduce interference within the selected 360 kHz-spaced narrow harmonic bands, and whether the remaining clean bands still preserve separable gesture patterns. Accordingly, this section proceeds from controlled body-coupling validation, to wideband harmonic selection, human-coupling robustness, vehicle-state robustness, and gesture separability.

We first compare spectra across four conditions: (1) no body contact with conductive cloth; (2) body contact, charger off; (3) body contact, charger on, no motion; and (4) body contact, charger on, hand near charging

region. This separates noise-floor increases from body contact versus peak enhancements from hand proximity, confirming that the body receives and modulates leakage components.

We then perform wideband scans under stable charging with body contact and hand near the charging region, repeating across multiple Tx/Rx combinations to verify consistency and guide selection of f_c.

Figure 2b shows the four-condition comparison: body contact raises the noise floor, while hand proximity produces pronounced peak enhancements. Figure 3a shows wideband spectral fingerprints; Figure 3b provides a 6–8 MHz enlargement. Observable structure appears near $f_0 \approx 360$ kHz multiples, with ∼7 MHz components exhibiting higher SNR. Based on this validation, we prioritize narrowband components in this region for gesture recognition.

We next examine whether the body-coupled channel remains observable under common human-factor variations. In this experiment, the conductive fabric is placed on the seat as the body-coupled electrode. Participants sit on the conductive fabric under different conditions and place their hand near the wireless charger, so that the body, seat electrode, and charger form the intended coupling path. We compare representative variations including upright sitting, offset sitting, different sitting/hand-placement actions, wearing outerwear, and different users. These factors may change the effective coupling coefficient among the body, conductive fabric, and charging pad, and therefore provide a direct test of whether the sensing baseline is overly sensitive to posture, clothing, or user identity.

Figure 4 shows that these human-factor variations slightly change the overall amplitude and baseline coupling strength, but the effect is not significant enough to obscure the useful harmonic structure. Across different actions, clothing conditions, and users, the 360 kHz-spaced comb-like components remain visible, and gesture-related spectral peaks are still observable on the selected harmonics. This indicates that human-dependent differences mainly appear as slow baseline shifts rather than a loss of the sensing channel. In practice, such baseline shifts can be compensated by the persistent calibration profile and runtime background update described in Sec. 5.5.

Since the target application is in-cabin interaction, we further evaluate whether representative vehicle operating states disturb the selected harmonic carriers. We consider four conditions shown in Fig. 5: engine start, low-speed driving, mild acceleration, and vehicle charging. These conditions introduce different types of disturbances, including engine-related electrical noise, motion-induced baseline fluctuation, and charging-related electromagnetic interference. This experiment therefore tests not only whether the baseline changes, but also whether such changes fall inside the narrow harmonic bands used by *MagHarp* .

Figure 5 shows that engine start, low-speed driving, and mild acceleration only introduce small baseline changes. Most of their disturbance energy is either broadband or located outside the selected narrow bands centered at integer multiples of 360 kHz, and can therefore be suppressed by the harmonic-band selection stage.

Vehicle charging produces stronger interference because the host vehicle's charging subsystem introduces additional high-power switching sources and coupling paths beyond those present during ordinary driving. Modern EV charging systems typically include on-board chargers, inverters, DC–DC stages, and high-voltage battery interfaces. These power-electronics modules generate fast voltage and current transients, and their parasitic capacitances, cable harnesses, DC supply lines, and common-mode/differential-mode current paths can convert conducted switching noise into field-bound electromagnetic emissions [25]. Prior EMC studies of EV and inductive automotive charging systems have shown that charging-related interference can appear over the 9 kHz–30 MHz range [32] and that the radiated magnetic field can be strongly affected by common-mode currents, DC supply-line layout, and stray capacitance. Therefore, unlike engine start or mild vehicle motion, which mainly changes the sensing baseline, vehicle charging can inject structured electromagnetic interference into a subset of the narrow harmonic neighborhoods used by *MagHarp* . This explains why several selected bands become contaminated in Fig. 5. Nevertheless, the contamination is band-selective rather than global: many harmonics remain sufficiently clean, allowing harmonic-band selection and stitching to discard unstable components while preserving gesture-discriminative information.

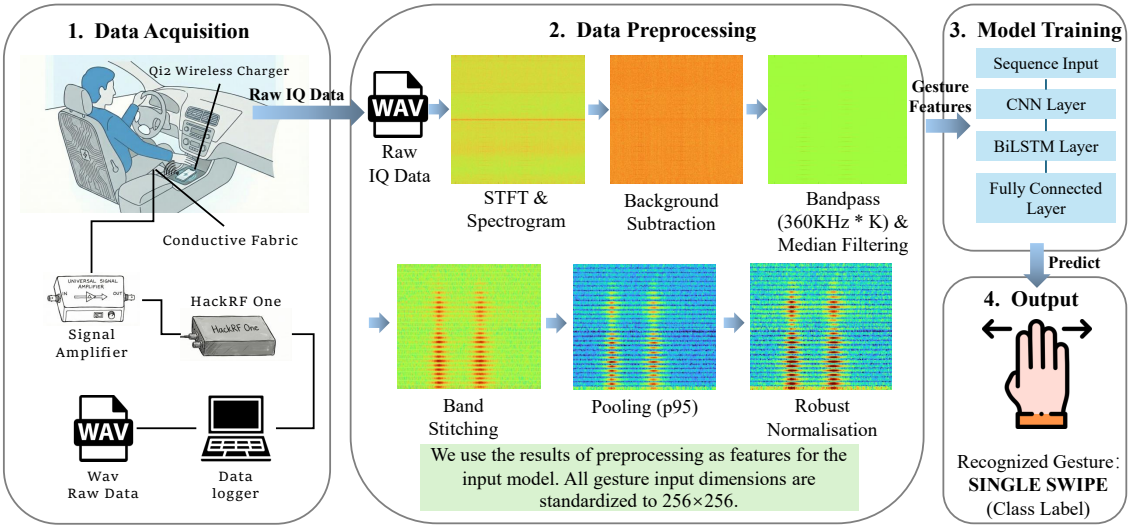


Fig. 7. System Overview of *MagHarp* . The pipeline captures raw electromagnetic leakage signals, applies physics-informed preprocessing (including STFT, background subtraction, and harmonic stitching) to extract robust features, and employs a hybrid CNN-BiLSTM network for real-time gesture classification. For deployment, MagHarp can optionally use the calibration and runtime adaptation mechanism described in Sec. 5.5.

After confirming that the body-coupled harmonic carriers remain observable under human-factor and vehicle-state variations, we finally examine whether different gestures produce distinguishable temporal patterns on these carriers. We study seven gesture classes in this work: swipe up, swipe down, single swipe, double swipe, single press, double press, and long press (illustrated in Fig. 11). For clarity, the color scale in figure is a robust-normalized intensity scale in [0,1], consistent with the feature construction in Sec. 5.3.4, where values are scaled by the 99.5th percentile of non-zero entries and clipped to [0,1]. Figure 6 shows seven gesture examples exhibiting distinct energy distributions, durations, and spectro-temporal shapes. These interpretable signatures provide a physical basis for our classifier (Section 5).

5 MAGHARP SYSTEM DESIGN

Guided by the physical model established and verified in §3 and §4, we present the design of *MagHarp* . The system aims to extract subtle hand-induced impedance variations from the noisy, high-energy background of wireless charging. As illustrated in Fig. 7, our pipeline consists of three stages: (i) physics-informed signal conditioning to isolate harmonic carriers, (ii) harmonic-band stitching for feature construction, and (iii) deep temporal classification for robust gesture decoding.

During practical deployment, *MagHarp* additionally supports an initial calibration and runtime adaptation mechanism, detailed at the end of this section. This mechanism builds a persistent user/vehicle profile once during setup or enrollment, reuses it across later sessions, and updates the background estimate automatically during operation.

5.1 Hardware Setup and Signal Acquisition

To convert the raw IQ waveforms captured by the HackRF into a stable and well-aligned representation for gesture recognition, we design a preprocessing and feature extraction pipeline based on *harmonic-band stitching*. The key idea is to suppress environmental/background interference and non-informative spectral regions in the

time–frequency domain, retain only the narrowband components around selected higher-order harmonics of the Qi2 fundamental (360 kHz), and stitch these bands into a fixed-size single-channel feature map for downstream classifiers.

Each gesture sample is stored as a two-channel IQ `.wav` file, where the I and Q streams are recorded in two separate channels. We form the complex baseband signal:

$$x[n] = I[n] + jQ[n]. \tag{7}$$

We compute a short-time Fourier transform (STFT) using a window length $w_{\text{len}} = 4096$, 50% overlap, and FFT size $n_{\text{fft}} = 8192$. Given the complex spectrogram $S(f, t)$, we convert magnitude to a dB scale:

$$P_{\text{dB}}(f, t) = 20 \log_{10} \left(|S(f, t)| + \epsilon \right), \tag{8}$$

where ϵ is a small constant for numerical stability. To reduce the influence of extremely weak components, we apply a dynamic-range truncation and keep only values within 80 dB of the global maximum.

5.2 Physics-Informed Signal Conditioning

5.2.1 Environmental Spectrum Estimation and Subtraction. To remove stable spectral patterns introduced by the environment and the sensing chain, we additionally record a background (no-gesture) IQ waveform and use it to construct an environmental spectral baseline (environment floor). Specifically, we compute $P^{\text{bg}}_{\text{dB}}(f, t)$ for the background waveform using the same STFT settings, and take the temporal median for each frequency bin:

$$\text{envFloor}(f) = \text{median}_t \left(P^{\text{bg}}_{\text{dB}}(f, t) \right). \tag{9}$$

For each gesture sample, we then perform background subtraction:

$$P_{\text{sub}}(f, t) = P_{\text{dB}}(f, t) - \text{envFloor}(f), \tag{10}$$

which emphasizes relative variations induced by the charging system and human-body coupling. In controlled experiments, envFloor is estimated from a separately recorded no-gesture reference. In deployment, this background estimate is provided and maintained by the calibration mechanism described in Sec. 5.5.

5.2.2 Harmonic Narrowband Selection (360 kHz-Spaced Harmonic Bands). Our measurements indicate that Qi protocols (e.g., Qi2) exhibit a relatively stable fundamental component around 360 kHz, and its higher-order harmonics appear as a set of observable narrowband spectral lines at higher frequencies. To exploit these stable lines as opportunistic carriers and improve cross-session consistency, we retain only the energy around integer-multiple harmonics of 360 kHz.

Let f_c denote the center frequency of a recording and f denote the STFT frequency axis. The absolute frequency is:

$$f_{\text{abs}} = f + f_c. \tag{11}$$

For each frequency bin, we compute the nearest harmonic center:

$$f_k = \text{round} \left(\frac{f_{\text{abs}}}{f_0} \right) f_0, \quad f_0 = 360\,\text{kHz}, \tag{12}$$

and define a keep mask with half-bandwidth $\Delta f = 5$ kHz:

$$\mathcal{M}(f) = \mathbb{I} \left(|f_{\text{abs}} - f_k| \le \Delta f \right). \tag{13}$$

All non-target regions are set to zero:

$$P_{\text{sub}}(f, t) \leftarrow P_{\text{sub}}(f, t) \cdot \mathcal{M}(f). \tag{14}$$

5.3 Harmonic-Band Stitching and Feature Construction

5.3.1 Time–Frequency Denoising and Smoothing. To further suppress residual narrowband interference and isolated outliers, we apply a 2D median filter (kernel size 3×3) to P_{sub}, yielding P_{med}. Since median filtering can slightly spread energy across neighboring frequency bins, we apply the mask $\mathcal{M}(f)$ again after filtering to ensure that only the target harmonic neighborhoods remain.

5.3.2 Harmonic-Band Stitching. After masking, the non-zero regions on the frequency axis consist of multiple disjoint harmonic sub-bands. To represent the joint information of multiple harmonic lines in a fixed-dimensional form, we extract the frequency bins belonging to each harmonic band ($f_k \pm \Delta f$) in ascending harmonic order and concatenate them along the frequency dimension, forming a new time–frequency matrix P_{stitch}:

$$P_{\text{stitch}} = \left[P_{k_{\min}};\, P_{k_{\min}+1};\, \cdots;\, P_{k_{\max}} \right], \tag{15}$$

where P_k denotes the spectral-energy block (all bins within the k-th harmonic band) over time. Note that the vertical axis of P_{stitch} no longer corresponds to a continuous frequency scale; instead, it represents an ordered harmonic-band sequence.

5.3.3 Pooling to a Fixed Size and Single-Channel Feature Map Construction. Because the number of STFT frames can vary across samples and to reduce downstream computational cost, we map P_{stitch} to a fixed size $H \times W$ using 2D pooling (we use 256×256). Concretely, we partition P_{stitch} into an $H \times W$ grid and aggregate values within each block. By default, we use 95th-percentile pooling (p95 pooling):

$$P_{\text{pool}}(i, j) = \text{percentile}\left(\mathcal{B}_{ij}, 95\right), \tag{16}$$

where $\mathcal{B}_{ij}$ denotes the set of elements within the (i, j)-th block. Compared to mean pooling, p95 pooling is more sensitive to salient local energy while remaining robust to a small number of extreme peaks.

5.3.4 Robust Normalization. To mitigate amplitude-scale variations across samples caused by factors such as device type and charging state (e.g., battery level), and to improve training stability, we apply robust normalization to P_{pool} and constrain the output to $[0, 1]$. Let $\Omega = \{(i, j) \mid P_{\text{pool}}(i, j) \neq 0\}$ be the set of non-zero locations. We use the 99.5th percentile of the non-zero values as a scale factor:

$$h = \text{percentile}\left(\left\{P_{\text{pool}}(i, j) \mid (i, j) \in \Omega\right\}, 99.5\right). \tag{17}$$

The normalized feature map is:

$$X(i, j) = \text{clip}\left(\frac{P_{\text{pool}}(i, j)}{h},\, 0,\, 1\right), \tag{18}$$

where $\text{clip}(\cdot)$ truncates values to $[0, 1]$. Each sample is finally stored as a single-channel feature tensor:

$$\mathbf{X}_1 \in \mathbb{R}^{256 \times 256 \times 1} \tag{19}$$

This representation preserves the spectro-temporal structure of harmonic lines while providing a unified input size suitable for both convolutional models and traditional feature pipelines.

5.3.5 Implementation and Output Organization. We batch-process all 7 gesture classes and save the resulting $\mathbf{X}_1$ as `.mat` files. The output directory is organized by gesture class to facilitate efficient loading during training and evaluation.

5.4 Lightweight Deep Temporal Gesture Classification

This section describes our training and classification pipeline for gesture recognition. Our primary objective is to learn a robust classifier that can withstand cross-day variations in conditions and remains resilient across different smartphones and wireless chargers.

5.4.1 Input Representation: From Feature Map to Sequence. After preprocessing, each sample is a 256×256 time–frequency feature map. To model temporal dynamics using 1D temporal convolutions and recurrent networks, we interpret the map as a length-256 sequence: the time axis corresponds to time steps, and each time step is a 256-dimensional frequency feature vector. Concretely, let $X \in \mathbb{R}^{256\times256}$ denote the feature map; the t-th column is treated as the input vector $\mathbf{x}_t \in \mathbb{R}^{256}$ at time step t.

Moreover, to reduce the dominance of absolute energy differences across frequency bins, we apply a row-wise z-score normalization along the time dimension:

$$\tilde{X}(f,t) = \frac{X(f,t) - \mu_f}{\sigma_f + \epsilon}, \quad \mu_f = \frac{1}{T}\sum_{t=1}^{T} X(f,t), \quad \sigma_f = \sqrt{\frac{1}{T}\sum_{t=1}^{T}\big(X(f,t) - \mu_f\big)^2}, \tag{20}$$

where $T = 256$ and ϵ is a small constant to avoid division by zero. The resulting sequence input has shape $256(\text{features}) \times 256(\text{timesteps})$.

5.4.2 Model Architecture: A CNN+BiLSTM Temporal Classifier. We employ a sequence classification network that combines local temporal pattern extraction with long-range dependency modeling. Fig. 8 illustrates our model architecture. The model takes a 256-dimensional sequence as input and does not apply additional normalization within the input layer, since row-wise standardization is performed during preprocessing. To capture local temporal patterns along the time axis, the network stacks two 1D convolutional blocks with kernel size 5 and 64/128 filters, respectively. Each convolution is followed by Batch Normalization and a ReLU activation, and a 1D max-pooling layer (pool size 2, stride 2) is used to downsample the temporal resolution and suppress high-frequency noise. Dropout with rates 0.05 and 0.15 is applied after the first and second convolutional blocks to mitigate overfitting. For global temporal modeling, a bidirectional LSTM (BiLSTM) layer with 128 hidden units is adopted with OutputMode=last to aggregate information over the entire sequence, followed by an additional Dropout layer with rate 0.25. Finally, a fully connected classification head maps the learned representation to the gesture categories, and the model outputs class probabilities through a Softmax layer and is trained end-to-end using cross-entropy loss over 7 classes.

5.4.3 Data Augmentation. To improve cross-day generalization, we apply data augmentation *only* to the training set, while keeping the validation and test sets unaugmented for a clean evaluation. Augmentations are performed on the time–frequency feature maps and include random time-scale perturbation to simulate variations in gesture speed with background-aware padding, time/frequency shifting with realistic padding strategies (edge replication for time and row-wise background filling for frequency), gamma/contrast adjustment and global gain jittering to model signal-strength fluctuations, as well as light noise injection and sparse pixel dropout to increase robustness to local corruption. After augmentation, all values are clipped to $[0, 1]$ to keep the input distribution well-formed.

5.4.4 Training Strategy and Model Selection. We train the network using the Adam optimizer with an initial learning rate of 3×10^{-4}, a maximum of 40 epochs, and a mini-batch size of 32, shuffling the training data every epoch. For each fold, the validation set is constructed from non-test days with a ratio of 0.15, while the test set is strictly held out by day to prevent temporal leakage. To reduce overfitting, we apply L_2 regularization with weight decay 1×10^{-2} and use dropout at multiple stages of the network.

Model selection is performed using validation monitoring with early stopping. Specifically, we evaluate on the validation set once per epoch (by setting the validation frequency to the number of iterations per epoch) and employ a patience of 8. We finally return and save the network parameters that achieve the lowest validation loss rather than the final-epoch weights, which improves robustness against occasional overfitting. In addition, we incorporate class-weighted cross-entropy by computing per-fold class weights from the training split, helping mitigate class imbalance during optimization.

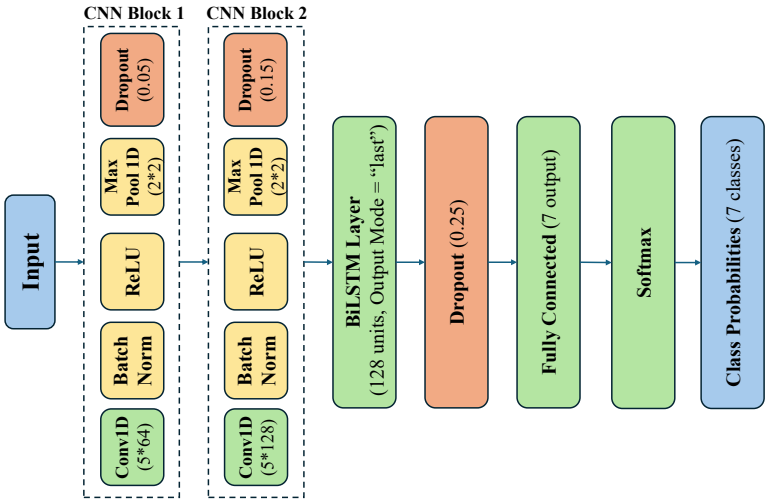


Fig. 8. Architecture of *MagHarp* gesture classifier. The network consists of two 1D convolutional blocks (Conv1D + BatchNorm + ReLU + MaxPool) with dropout regularization, followed by a BiLSTM layer (128 units, OutputMode=`last`), a dropout layer, and a fully connected layer with softmax to output probabilities over 7 gesture classes.

5.4.5 Inference and Evaluation Metrics. After training each fold, we evaluate the model on both the validation set and the independent test day. For each test sample, the model outputs a predicted label $\hat{y}$, and we compute classification accuracy by comparing it with the ground-truth label y:

$$\text{Acc} = \frac{1}{N}\sum_{i=1}^{N}\mathbb{I}\left(\hat{y}_i = y_i\right). \tag{21}$$

We also report confusion matrices to analyze dominant confusion patterns among gestures. Finally, by evaluating the mean accuracy and F1 score across folds, we summarize the cross-day generalization performance.

5.5 Initial Calibration and Runtime Adaptation

While the generic pipeline described above already provides robust gesture recognition, practical deployment introduces additional variability from ambient electromagnetic interference, charger operating state, body–electrode contact condition, clothing thickness, vehicle-induced coupling changes, and individual gesture habits. To account for these factors without retraining the full network, *MagHarp* supports an initial calibration and runtime adaptation mechanism. Importantly, this mechanism is not intended to be repeated before every use. Instead, *MagHarp* builds a persistent deployment-specific calibration profile during initial setup, user enrollment, or when a substantially new user/vehicle condition is introduced. The resulting profile can be stored and reused across subsequent driving sessions, while the background estimate is automatically updated during operation using high-confidence gesture-free intervals. The mechanism consists of two components: (i) initial electromagnetic background calibration with runtime background tracking for stable feature extraction, and (ii) optional lightweight network adaptation for larger user- or condition-specific distribution shifts.

5.5.1 Initial Electromagnetic Calibration and Runtime Background Update. During initial setup or enrollment, the target user sits in a natural driving posture with the phone placed on the charging pad and remains still for a short gesture-free interval. We use this short clip to estimate an initial deployment-specific environmental floor, following the same STFT, log-magnitude conversion, and temporal-median procedure described in Sec. 5.1. This

initial calibration captures not only ambient electromagnetic clutter and sensing-chain background, but also the current charger/phone operating state and the user's static coupling condition, which may vary with body shape, clothing material, seating posture, and contact quality with the conductive fabric. The estimated floor is stored as part of the user/vehicle calibration profile and reused in later sessions.

Let p denote the stored deployment profile. We replace the generic background floor with the profile-specific floor $\mathrm{envFloor}_p(f)$ and perform calibrated background subtraction:

$$P_{\mathrm{sub}}(f,t) = P_{\mathrm{dB}}(f,t) - \mathrm{envFloor}_p(f). \tag{22}$$

This profile-specific subtraction suppresses stationary components associated with the current vehicle, charger, phone, environment, and static body coupling, while preserving dynamic perturbations caused by hand motion.

During online operation, *MagHarp* further maintains this profile without asking the user to repeat manual calibration. When the system detects a high-confidence gesture-free window, the environmental floor can be updated using a slow temporal-median buffer or an exponential moving average:

$$\mathrm{envFloor}_p^{(m)}(f) = (1-\lambda)\mathrm{envFloor}_p^{(m-1)}(f) + \lambda\,\mathrm{median}_t\left(P_{\mathrm{dB}}^{(m)}(f,t)\right), \tag{23}$$

where m indexes the accepted gesture-free update window and λ is a small update rate. The update is confidence-gated so that active gesture windows are not absorbed into the background model. This runtime update tracks gradual changes caused by charging-state drift, posture adjustment, temperature, or clothing/contact variation, making calibration a persistent and self-updating profile rather than a per-use manual procedure.

5.5.2 Optional Lightweight Network Adaptation. After profile-specific background calibration, residual distribution shifts may still remain. For example, different users may perform gestures with different speeds, amplitudes, trajectories, and execution rhythms, while in-vehicle motion or charging-state changes may alter the local spectro-temporal patterns of the stitched harmonic features. To reduce this mismatch, *MagHarp* optionally supports lightweight network adaptation using a small labeled adaptation subset collected under the target user or operating condition.

Let

$$\mathcal{D}_p^{\mathrm{adapt}} = \{(\mathbf{X}_i, y_i)\}_{i=1}^{N_p} \tag{24}$$

denote the adaptation set associated with deployment profile p, where $\mathbf{X}_i$ is the calibrated and stitched feature map and y_i is the corresponding gesture label. In our evaluation, we intentionally use 5 labeled samples per gesture class as a fixed stress-test adaptation protocol. This setting is used to demonstrate that the model can recover from large user- or condition-specific domain shifts with limited labeled data; it should not be interpreted as a required per-use calibration burden. In practical deployment, the adapted profile can be learned once and stored, and subsequent sessions can reuse this profile together with the runtime background update described above. The required labeled effort can be further reduced by using fewer representative gestures, confidence-based pseudo-labeling, or a pre-collected library of user/vehicle/charging-state feature profiles.

Starting from the pretrained generic model with parameters θ_0, we do not retrain the entire CNN-BiLSTM network. Instead, we freeze the early low-level feature extractor and the recurrent temporal module, including `conv1`, `bn1`, and the BiLSTM layer, and only update a small set of higher-level adaptation layers: `conv2`, `bn2`, and the final fully connected layer. This design allows the model to adjust to local domain-specific feature shifts while preserving the generic temporal representation learned from the main training data.

Let $\theta_{\mathrm{adapt}} = \{\theta_{\mathrm{conv2}}, \theta_{\mathrm{bn2}}, \theta_{\mathrm{fc}}\}$ denote the trainable parameters during adaptation. The adapted model is obtained by optimizing

$$\theta_p^* = \arg\min_{\theta_{\mathrm{adapt}}} \sum_{(\mathbf{X}_i, y_i)\in\mathcal{D}_p^{\mathrm{adapt}}} \mathcal{L}\left(f(\mathbf{X}_i; \theta_0 \setminus \theta_{\mathrm{adapt}}, \theta_{\mathrm{adapt}}), y_i\right), \tag{25}$$

where $\mathcal{L}$ is the cross-entropy loss and all parameters outside θ_{adapt} remain fixed.

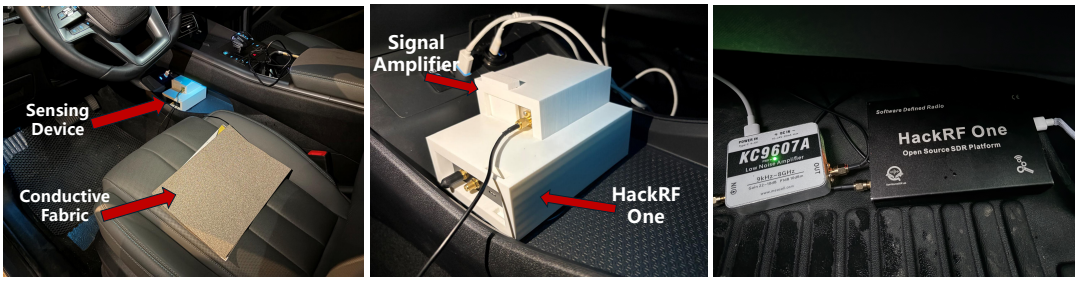
In our evaluation implementation, adaptation uses the Adam optimizer with a mini-batch size of 4, a learning rate of 5×10^{-4}, and 10 epochs. We apply L_2 regularization with weight decay 1×10^{-2} and do not apply data augmentation during adaptation. To make the classifier adjust faster than the feature layers, we use a larger learning-rate factor for the final fully connected layer than for `conv2` and `bn2`. Specifically, `conv2` and `bn2` use a learning-rate factor of 1, while the final fully connected layer uses a factor of 5. This keeps the adaptation lightweight while allowing both the higher-level local feature extractor and the decision boundary to shift toward the current user or operating condition.

At deployment time, incoming signals are first processed using the stored environmental floor $\text{envFloor}_p(f)$, which is continuously maintained through gesture-free runtime updates. The resulting feature maps are then fed into either the generic model or the stored adapted model, depending on whether optional labeled adaptation has been enabled for the current profile. This design allows *MagHarp* to support two practical modes: a zero-label mode that only uses persistent background calibration and runtime update, and an adapted mode that uses a small one-time enrollment/adaptation set to update `conv2`, `bn2`, and the final fully connected layer.

6 IMPLEMENTATION

We implemented a data-acquisition prototype with offline processing of *MagHarp* to evaluate its performance in realistic driving scenarios. This section details the hardware deployment, the data acquisition protocol, and the dataset composition used for evaluation.

6.1 Hardware Prototype



(a) Actual deployment scenario. (b) Example in-vehicle placement of the integrated module. (c) Signal amplifier and HackRF.

Fig. 9. *MagHarp* hardware setup and signal acquisition. We place the conductive fabric on the seat. The signal amplifier and HackRF are integrated into a compact enclosure for convenient deployment, which can be placed in the gap beneath the center console or underneath the seat.

As shown in Fig. 9, we deploy a piece of flexible conductive fabric on the seat surface as a body-coupled sensing electrode to capture wireless-charging leaked signals through the human body. The fabric electrode has a size of 20×30 cm and costs approximately $2. The electrode is connected via an RF coaxial cable to a low-frequency/wideband signal amplifier (KC9607A), whose effective operating range covers 9 kHz–8 GHz. The amplifier output is then fed into a software-defined radio receiver (HackRF One) for complex baseband (I/Q) sampling. For practical deployment, we further integrate the signal amplifier and HackRF into a compact enclosure, which can be conveniently placed in the gap beneath the center console or underneath the seat. We defer a broader discussion of miniaturization and embedded automotive integration to sec. 8.3.



Fig. 10. Cars equipped with wireless charging pads.

We configure the HackRF One with a sampling rate of 10 MS/s, a center frequency of $f_c = 10$ MHz, and a receive bandwidth of $B = 10$ MHz. This setup is intended to capture the high-frequency harmonic components of the 360 kHz wireless-charging signal in the several-MHz range (with prominent energy around ~ 7 MHz), while avoiding placing the strongest component exactly at the center of the SDR passband to reduce potential DC/center-bin artifacts and noise interference. We collect data using SDRSharp. During acquisition, we disable the built-in noise suppression functions and turn off AGC (automatic gain control) to ensure consistent and reproducible recordings. We fix the gain settings with both VLA Gain and VGA Gain set to 16 dB, which effectively enhances the gesture-induced signal variations in practice. The recorded raw I/Q streams are saved for offline processing.

We implement the signal processing pipeline proposed in Sec. 5 in MATLAB. In general, the offline processing first converts the recorded I/Q samples into an amplitude/envelope representation, and then applies the proposed denoising, normalization, and feature construction steps (e.g., time–frequency analysis and background suppression). Finally, the extracted features are fed into the gesture classifier for recognition. All offline processing is performed on a desktop PC equipped with an Intel Ultra 5 245KF CPU and an NVIDIA 5060 Ti GPU with 16 GB memory.

We believe this design can be well integrated into future automotive scenarios. The conductive fabric can be used directly as part of the seat upholstery or embedded beneath the seat surface to enable signal acquisition without affecting user comfort or driving experience; the amplifier and sampling module can be packaged into a compact enclosure and integrated into the vehicle interior (e.g., in the gap beneath the center console or underneath the seat). At the same time, we observe that the wireless chargers in newly launched vehicles are positioned precisely on the center console (as shown in Fig. 10), which is the prime user interaction zone. From a safety perspective, our system is purely passive and does not inject any additional excitation into the human body. The induced current caused by the wireless-charging leakage field is extremely small. Our experimental protocol was approved by the ethics review board of our university.

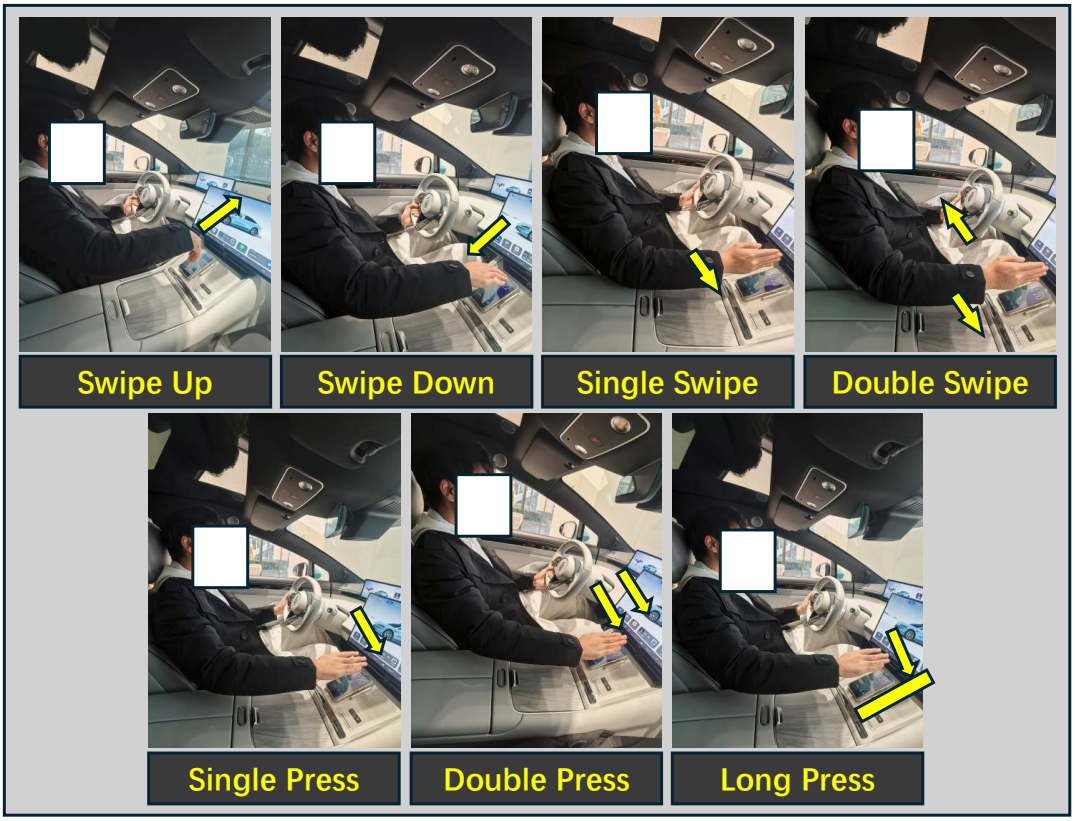


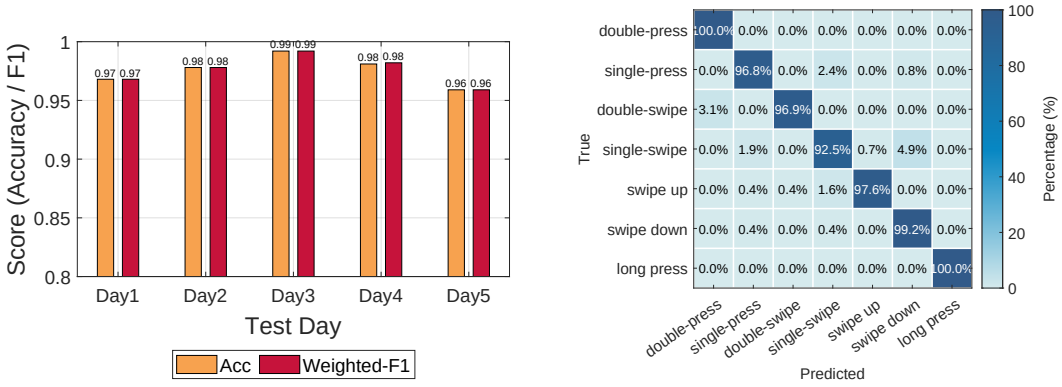
Fig. 11. Illustrations of the seven gestures performed in the vehicle. Swipe up, swipe down denote a swipe in the respective direction. Single swipe denotes swipe once away from the user's side. Double swipe denotes swipe away from the user first, then swipe back. Single press or double press signifies pressing the palm flat towards the device the corresponding number of times. Long press denotes pressing and holding for a duration.

6.2 Data Collection Protocol

Experiments were conducted using a commercial wireless charging pad (Belkin Boost Charge Pro 15W) with an iPhone 17 (iOS 26.2) as the receiver device (Rx) charging normally throughout data collection. The operator sat on a conductive fabric electrode that served as a body-coupled sensing interface. The parameters and settings for each accessory are described in Section 6. To evaluate robustness under realistic electromagnetic interference, experiments were conducted in a noisy laboratory environment without any electromagnetic shielding.

We collected seven gesture classes as shown in Fig. 11: *swipe up*, *swipe down*, *single swipe*, *double swipe* (each a single swipe), *single press*, *double press*, *long press*.

Data were collected over 10 days. Each day, we recorded approximately 50 instances per gesture class at a gesture distance of 15 cm, yielding about 350 samples per day and approximately 3,500 samples in total. A primary operator (male, 23 years old) performed all gestures following the gesture definitions described in Fig. 11. To assess cross-user variability, we additionally recruited 10 volunteers to collect an auxiliary dataset under the same setup and protocol at the same distance, where each volunteer performed approximately 20 instances per gesture class, resulting in 1,400 additional samples (10×7×20). This auxiliary dataset was used only for cross-user evaluation (not included in training). To reduce charging-state bias, the phone battery state-of-charge was kept within



(a) Accuracy and weighted F1 score bar chart for each fold in 5-fold cross-validation. As shown in the figure, the accuracy exceeded 95% on all test days.

(b) Confusion matrix for all test data. We observe a clear concentration along the diagonal.

Fig. 12. 5-fold Leave-One-Day-Out evaluation results chart. (a) daily accuracy rate and F1 score, (b) total confusion matrix.

20%–80% during collection. After acquisition, we followed the preprocessing pipeline described in Section 5.1. The corresponding dataset has been open-sourced at GitHub[1].

We adopt a recording-day–based cross-validation strategy to prevent temporal leakage. Under the same acquisition protocol and device configuration, we select five representative recording days that are widely spaced in time and perform 5-fold Leave-One-Day-Out (LODO) evaluation: in each fold, data from one day is used as an independent test set, while data from the remaining days is used for training and validation.

7 EVALUATION

We evaluate *MagHarp* across multiple dimensions: (1) overall recognition accuracy and cross-day generalization; (2) the contribution of each algorithmic component; (3) robustness under realistic operating factors, including in-vehicle dynamics, device diversity, charger models, battery charging states, and user variability; and (4) the usable interaction distance. We further examine a phone-absence setting to explore the feasibility of short-range standby interaction without an active receiver.

7.1 Evaluation Metric

We report overall accuracy as the primary metric. In addition, we provide per-class precision and F1-score, together with the confusion matrix for error pattern analysis.

Let $\mathbf{M} \in \mathbb{N}^{K\times K}$ denote the confusion matrix with $M_{i,j} = |\{n : y_n = i,\ \hat{y}_n = j\}|$ (row: ground-truth, column: prediction). For each class k, define $\mathrm{TP}_k = M_{k,k}$, $\mathrm{FP}_k = \sum_{i\neq k} M_{i,k}$, and $\mathrm{FN}_k = \sum_{j\neq k} M_{k,j}$. We compute:

$$P_k = \frac{\mathrm{TP}_k}{\mathrm{TP}_k + \mathrm{FP}_k}, \qquad F1_k = \frac{2\,\mathrm{TP}_k}{2\,\mathrm{TP}_k + \mathrm{FP}_k + \mathrm{FN}_k}, \tag{26}$$

where we set the metric to 0 if the denominator is zero.

7.2 Overall Performance and Cross-Day Generalization

Overall performance. Fig. 12 shows strong cross-day generalization in 5-fold LODO evaluation: both accuracy and weighted-F1 remain within 95.9–99.2%, with a mean accuracy of 97.6%. The confusion matrix is highly diagonal, indicating stable separability across all seven gestures.

The errors concentrate among swipe-like classes with similar temporal structures: *single-swipe* has the lowest recall (92.5%) and is mainly confused with *swipe down* (4.9%) and *single-press* (1.9%), while *double-swipe* is occasionally predicted as *double-press* (3.1%). Overall, most confusions occur within gestures that share similar motion direction or temporal/repetition patterns. These actions are inherently quite similar, so distinguishing between them can be somewhat challenging.

Comparison with Baselines. For a fair comparison, all baseline models were evaluated under the same 5-fold representative-day leave-one-day-out protocol as *MagHarp* . XGBoost, LightGBM, AutoGluon, and TabPFN take the same stitched time-frequency representation as input after flattening, with PCA fitted only on the training split of each fold and then applied to the corresponding validation and test sets. LSTM receives the same representation as a temporal sequence, while 2D-CNN uses the same 2D map directly. Hyperparameters are selected on the same validation split (15% of the non-test days), and final results are reported on the held-out test day.

Figure 13 compares *MagHarp* with a diverse set of baseline classifiers, including conventional machine-learning methods, tabular learners, automated ensemble methods, and deep neural models. Overall, *MagHarp* achieves the best performance, reaching 0.97 accuracy and 0.97 F1.

Among all baselines, TabPFN is the strongest competitor, achieving 0.87 accuracy and 0.86 F1. LightGBM follows with 0.83 accuracy and 0.82 F1, while XGBoost and AutoGluon yield comparable results at around 0.81–0.83 accuracy and 0.81–0.82 F1. In contrast, the deep baselines remain noticeably weaker: LSTM attains 0.78 accuracy and 0.75 F1, and 2D-CNN further drops to 0.61 accuracy and 0.59 F1. Compared with the strongest baseline, *MagHarp* still preserves an advantage of approximately 10 percentage points in both accuracy and F1.

These comparisons suggest that the advantage of *MagHarp* is not merely due to using a larger model. Strong tabular learners such as TabPFN, LightGBM, XGBoost, and AutoGluon already perform competitively on the same stitched representation, indicating that the signal pipeline itself extracts informative features. However, the consistent margin of the CNN-BiLSTM suggests that explicitly modeling local spectro-temporal structure together with longer-range temporal dependencies remains beneficial for this task, which justifies our final model choice.

7.3 Ablation Study

To quantify the contribution of each design component in *MagHarp* , we conduct a progressive ablation study. Starting from the full system, we remove key signal pre-processing modules (background subtraction, 2D median filtering, and harmonic sub-band stitching), architectural components (CNN-based feature extraction and BiLSTM-based temporal modeling), as well as the training strategy (data augmentation). We report test-set accuracy as the evaluation metric.

Table 2 quantifies the contribution of each component in MagHarp. The full system achieves 97.6% accuracy on the test set. Removing key preprocessing modules leads to clear performance degradation: without background subtraction, the accuracy drops to 85.4% (−12.2%), indicating that suppressing the static environmental baseline is essential for exposing subtle gesture-induced perturbations. Likewise, removing harmonic sub-band stitching reduces the accuracy to 79.3% (−18.3%), and removing 2D median filtering results in 80.0% (−17.6%), showing that both multi-band fusion and robust denoising are critical for stable signal representation.

[1] https://github.com/MagHarp/MagHarp

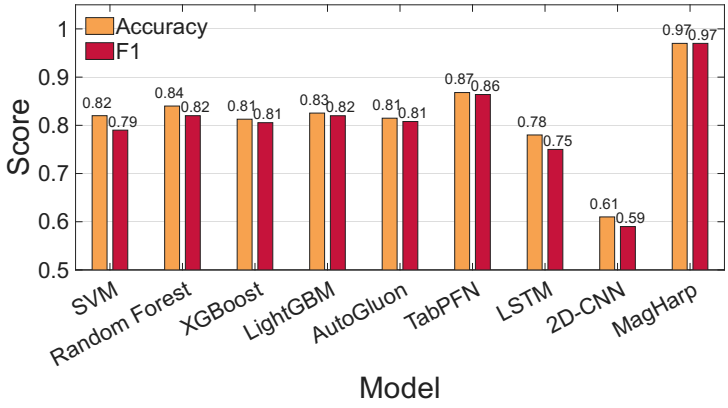


Fig. 13. Performance comparison between *MagHarp* and representative baseline models. *MagHarp* achieves the highest accuracy and F1, consistently outperforming conventional machine-learning, tabular, ensemble, and deep-learning baselines.
Table 2. Ablation study quantifying the contribution of each design component. The test set accuracy drops significantly when key preprocessing or architectural modules are removed.

Method Configuration	**Accuracy (%)**	**Δ (%)**
Full System (MagHarp)	**97.6**	–
Signal Preprocessing		
w/o Background Subtraction	85.4	−12.2
w/o Harmonic Sub-band Stitching	79.3	−18.3
w/o 2D Median Filtering	80.0	−17.6
Model Architecture		
w/o BiLSTM Temporal Module	63.7	−33.9
w/o CNN Feature Extractor	85.7	−11.9
Training Strategy		
w/o Data Augmentation	96.2	−1.4

Among the architectural components, the BiLSTM temporal module contributes the most: removing it causes the largest drop, down to 63.7% (−33.9%), which highlights the importance of modeling temporal dependencies in gesture dynamics. Removing the CNN feature extractor also significantly lowers performance to 85.7% (−11.9%), confirming the need for effective local feature learning before temporal aggregation. In contrast, removing data augmentation yields a relatively small reduction to 96.2% (−1.4%), suggesting that augmentation provides additional robustness but is less critical than the core preprocessing and model design. These results demonstrate that MagHarp's performance relies on the joint effect of signal preprocessing, spatial-temporal feature extraction, and temporal modeling, with the temporal module being particularly important.

7.4 Robustness Across Real-World Variables

To evaluate *MagHarp* 's robustness against real-world deployment factors, we conduct a systematic set of experiments across in-vehicle operating conditions, smartphone models, wireless charger models, battery state-of-charge regimes, and user variability. It is worth noting that the robustness results in this section are evaluated on newly collected, independent test sets rather than on the original samples used in Sec. 7.2. All experimental

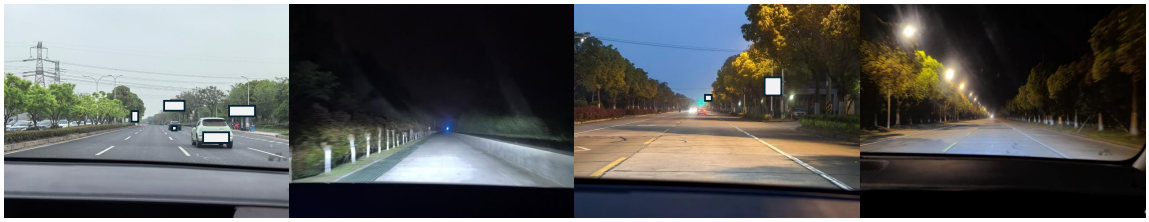


Fig. 14. Representative real-world driving scenarios used for on-road data collection: (a) daytime normal driving on public roads, (b) driving on rough roads to collect vibration-induced disturbance data, (c) driving on roads with frequent intersections to collect turning-related gesture data, and (d) nighttime acceleration on lightly trafficked roads. These scenarios capture typical in-cabin perturbations caused by routine driving, road-surface vibration, turning maneuvers, and acceleration.

settings remain the same as in the standard setting, except for the single factor under study in each subsection. Small performance differences are thus expected due to session-to-session variation in gesture execution, body coupling, and charging-state baseline.

7.4.1 In-Vehicle Operating Conditions. To directly evaluate robustness under realistic cabin dynamics and practical deployment conditions, we collect independent in-vehicle test data under five representative operating conditions: *parked EV charging*, *low-speed straight driving*, *mild acceleration*, *turning/body tilt*, and *rough-road vibration*. These conditions cover major factors that may affect in-cabin gesture recognition, including electromagnetic interference during EV charging, routine vehicle motion, acceleration-induced body-coupling variation, steering-induced body tilt, and road-surface vibration. Fig. 14 illustrates representative on-road data-collection scenarios, including normal driving, rough-road driving, turning-intensive roads, and nighttime acceleration.

For each operating condition, the reported results in Table 3 are computed on newly collected in-vehicle data. To evaluate whether *MagHarp* can recover from severe condition-specific domain shifts, we adopt a fixed stress-test adaptation protocol. For each operating condition, we select 5 samples per gesture class from the newly collected data, resulting in 35 labeled adaptation samples in total. The remaining samples are held out for testing. The *Before FT* result is obtained by directly applying the generic model to the held-out test samples. The *After FT* result is obtained after adapting the model with the 35 selected samples and evaluating on the same held-out test samples. No held-out test sample is used during adaptation. We emphasize that this protocol is designed for experimental recoverability analysis rather than to prescribe a per-use calibration requirement. In deployment, the resulting adaptation profile can be created once, stored, reused across sessions, and continuously maintained through runtime background updates.

Table 3 shows that the generic model experiences a clear domain shift when directly applied to newly collected in-vehicle data. Before fine-tuning, the accuracy ranges from 33.3% to 55.7%. The parked EV-charging condition achieves the highest before-FT accuracy of 55.7%, while dynamic driving conditions, including low-speed driving, acceleration, turning/body tilt, and rough-road vibration, are more challenging, with before-FT accuracy around 33.3–36.4%. This indicates that vehicle motion not only changes the body-coupling baseline and introduces motion-related disturbances, but may also slightly deform users' gesture execution during driving. For example, acceleration, turning, or road vibration can cause small deviations in hand trajectory, gesture amplitude, or timing, thereby shifting the feature distribution from the standard training condition.

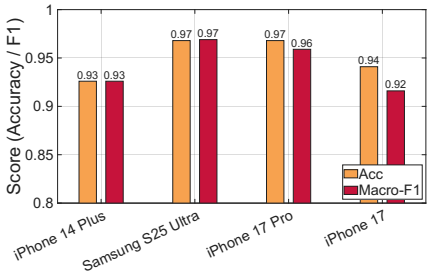
Under this stress-test adaptation protocol, the recognition accuracy improves substantially across all tested conditions. *MagHarp* reaches 81.0% under parked EV charging, 86.9% under low-speed straight driving, 84.8% under mild acceleration, 83.3% under turning/body tilt, and 83.0% under rough-road vibration. The largest gains appear in dynamic driving conditions, with improvements of 48.1–52.5 percentage points. These results suggest that in-vehicle disturbances mainly introduce recoverable condition-specific baseline shifts and coupling changes rather than eliminating gesture-discriminative information.

Table 3. Gesture recognition performance under representative in-vehicle operating conditions. FT uses 5 labeled samples per gesture class from the target condition.

Condition	Before FT Acc. (%)	After FT Acc. (%)	Δ (%)
Parked, EV charging	55.7	81.0	25.3
Low-speed straight driving	34.4	86.9	52.5
Mild acceleration	36.4	84.8	48.4
Turning / body tilt	33.3	83.3	50.0
Rough-road vibration	34.9	83.0	48.1



(a) Different phone models used in the experiment.



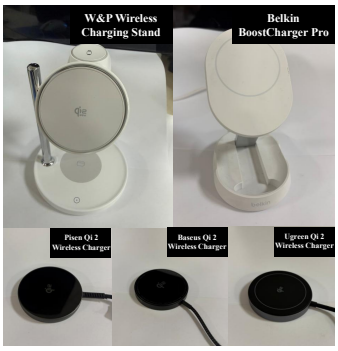
(b) Accuracy rate of our method tested on four mobile phones.

Fig. 15. Generalizability experiment diagram for different mobile phones.

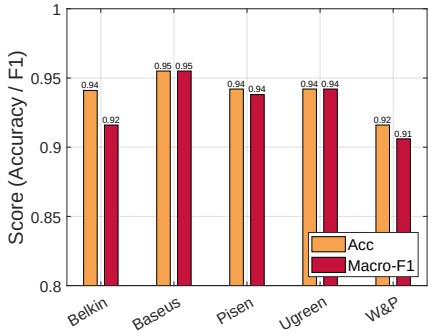
Among the tested conditions, low-speed straight driving achieves the highest after-FT accuracy, while parked EV charging and rough-road vibration remain relatively more challenging. This trend is consistent with our feasibility-stage spectral observations: EV charging can introduce stronger electromagnetic interference, whereas rough-road vibration causes stronger body-pressure redistribution and motion-induced baseline fluctuation. Nevertheless, after adaptation, all tested conditions remain above 81% accuracy, suggesting that the tested in-vehicle disturbances mainly introduce recoverable domain shifts rather than eliminating gesture-discriminative information.

7.4.2 Influence of Smartphone Diversity. To examine the impact of device hardware differences and vendor-specific protocol implementations, we test four representative smartphones: Samsung S25 Ultra, iPhone 14 Plus, iPhone 17, and iPhone 17 Pro as Fig. 15a shows. Results are in Fig. 15b. The results show that *MagHarp* maintains stable recognition performance across different phone models, indicating that the extracted charging-side signals and the downstream inference pipeline are insensitive to phone brand, internal structural variations, and implementation details, demonstrating strong cross-device robustness.

7.4.3 Influence of Charger Models. To evaluate whether charger-side design diversity affects the background spectrum and thus the system performance, we experiment with five commercially available wireless chargers from different manufacturers, including Belkin, Ugreen, W&P, Pisen, and Baseus as Fig. 16a shows. Since these products vary in transmitter-coil structure, control circuitry, and mechanical design, they may produce different harmonic



(a) Different models of wireless chargers used in the experiment.



(b) Accuracy rate of our method tested on five wireless chargers.

Fig. 16. Generalizability experiment diagram for different wireless chargers.

backgrounds and noise distributions. Results are shown in Fig. 16b. Our results demonstrate that *MagHarp* remains stable across all tested chargers, suggesting that the proposed signal processing pipeline effectively suppresses charger-related stable background components while preserving gesture-induced perturbation features, thereby achieving robustness to charger variations.

7.4.4 Dynamic Charging States. To directly evaluate robustness under a broader range of practical charging conditions, we additionally test *MagHarp* under three battery regimes: *low SoC* (< 20%), *mid SoC* (20%–80%), and *high SoC* (> 80%).

Table 4 summarizes gesture recognition performance under different battery state-of-charge (SoC) regimes. Without additional calibration, MagHarp achieves 90.3% accuracy at low SoC (< 20%), 97.6% in the standard mid-SoC range (20%–80%), and 86.7% at high SoC (> 80%). As expected, performance degrades outside the mid-SoC regime, with the most noticeable drop appearing at high SoC, where many phones enter the trickle-charging phase and the transmitted field becomes weaker. However, using the stored adaptation-profile mechanism improves the low-SoC accuracy to 93.3% (+3.0 points) and the high-SoC accuracy to 91.2% (+4.5 points), showing that deployment-specific adaptation can effectively mitigate charging-state-induced variability without requiring a new manual calibration for each charging state.

To further examine how charging-state dynamics affect the sensing baseline, we record a complete wireless charging process from 15% to 100% battery level and plot the charging-power trace in Fig. 17. The trace shows a clear multi-stage charging behavior. At the beginning, the charging power quickly stabilizes at around 18 W. As the phone temperature rises, the charger dynamically reduces the power; after the phone cools down, the power partially recovers, but this stage remains around 13 W overall. When the battery level further increases, the charging power enters a lower plateau of approximately 9 W. Finally, near the high-SoC regime, the phone enters a trickle-charging stage, where the power gradually decreases to around 2.5 W.

Importantly, Fig. 17 indicates that charging-state changes mainly occur as long-timescale stage-wise power adaptation rather than large instantaneous power jumps. The short-term power fluctuations within each stage are small compared with the duration of a typical gesture window, while the major power-level changes evolve over hundreds to thousands of seconds. Therefore, charging-state variation is more likely to introduce a slowly varying baseline shift than an abrupt gesture-like transient. This observation is consistent with the SoC robustness results in Table 4: although performance decreases in low-SoC and high-SoC regimes, the adaptation-profile mechanism improves the accuracy from 90.3% to 93.3% at low SoC and from 86.7% to 91.2% at high SoC. These results show

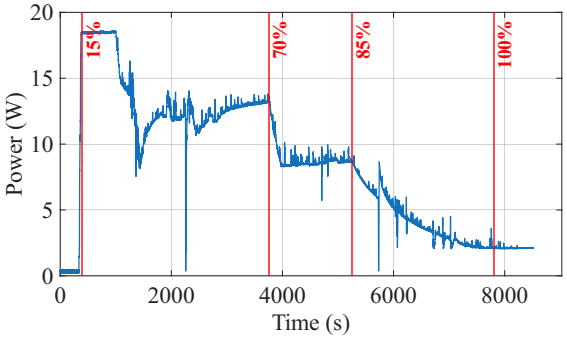


Fig. 17. Full charging-power trace from 15% to 100% battery level. The charging process exhibits four long-timescale stages: an initial high-power stage around 18 W, a thermally adjusted stage around 13 W, a lower-power plateau around 9 W, and a final trickle-charging stage that gradually decreases to around 2.5 W. These changes are mainly slow stage-wise adaptations rather than large instantaneous jumps, suggesting that charging dynamics primarily introduce baseline shifts rather than gesture-like transients.

Table 4. Gesture recognition performance under different battery state-of-charge (SoC) regimes before and after using the deployment-specific adaptation-profile mechanism.

Battery regime	**Accuracy (%)**	**Accuracy after adaptation (%)**	**Δ (%)**
Low SoC (< 20%)	90.3	93.3	3.0
Mid SoC (20–80%)	97.6	/	/
High SoC (> 80%)	86.7	91.2	4.5

that *MagHarp* remains robust to practical charging-state changes when using a deployment-specific adaptation profile. In deployment, gradual SoC-related baseline drift can be further tracked by the runtime background update described in Sec. 5.5, rather than requiring a new manual calibration for each charging state.

In the unlikely case of a short step-like power change within a gesture window, MagHarp does not rely on absolute signal amplitude alone. The background subtraction emphasizes relative perturbations, the narrow harmonic-band selection and 2D median filtering suppress isolated spectral outliers, and robust normalization reduces sample-level amplitude scaling caused by charging-state changes. The CNN-BiLSTM then classifies the remaining spectro-temporal gesture pattern rather than a single global power level, which helps avoid mistaking slow or step-like charging-state variations for gestures.

7.4.5 User Variability and Personalization. Different users naturally exhibit variations in hand size, gesture amplitude, execution speed, and habitual motion patterns, which may lead to distribution shifts between individuals. To assess robustness to inter-user variability under a controlled yet practical setting, we recruit ten volunteers with diverse anthropometric characteristics and daily interaction habits to participate in the experiments.

Before data collection, all participants are introduced to our *pre-defined standard gesture set* (i.e., the canonical definition of each gesture used as the system reference). We provide a short training session that includes (i) an explanation of the intended semantics and start/end poses of each gesture, (ii) a visual demonstration of the correct execution, and (iii) a guided practice phase. During this phase, participants confirm that they understand the definition of every gesture and can perform them consistently. This procedure helps ensure that performance differences mainly reflect user-dependent execution variability rather than misunderstanding of gesture meanings.

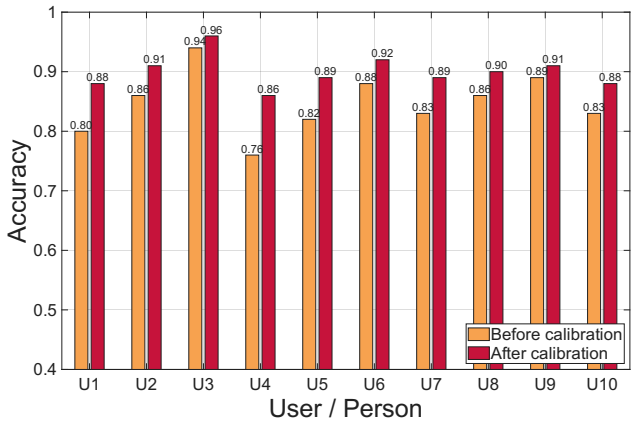


Fig. 18. Cross-user performance before and after building a user-specific adaptation profile.

For each user, we collect data by asking them to perform every gesture approximately 20 times following the same acquisition protocol. To approximate real usage while keeping conditions comparable across users, repetitions are recorded continuously with natural transitions, and participants are encouraged to execute gestures in their comfortable manner (e.g., their preferred speed and amplitude) as long as they adhere to the standard definition. Short breaks are provided to reduce fatigue effects and maintain execution quality.

We report two settings: before adaptation, where the generic model is directly tested on the target user, and after adaptation, where a user-specific lightweight adaptation profile is built following the mechanism described in Sec. 5.5. Results are shown in Fig. 18. Before adaptation, the generic model achieves an average accuracy of 84.7% across the 10 users, indicating that *MagHarp* can generalize to unseen users but is still affected by user-dependent coupling conditions and gesture execution styles. After adaptation, the average accuracy improves to 90.0%, yielding an overall gain of 5.3 percentage points. These results show that lightweight personalization can compensate for variations in gesture amplitude, execution rhythm, and body–electrode coupling.

This experiment is intended as an auxiliary cross-user transferability study. The main dataset is centered on one primary operator, and the multi-user dataset is used for held-out evaluation and lightweight personalization rather than for training a fully user-independent model. Therefore, the results should be interpreted as evidence of transferability and adaptation effectiveness, rather than as a complete user-independent benchmark.

7.5 Interaction Distance

We evaluate the usable interaction distance of *MagHarp* in the standard phone-present charging setting. During data collection, the user's hand is placed at different distances from the wireless charger, ranging from 0 cm to 30 cm with a 5 cm interval. For each distance, we compute the gesture-induced SNR over the selected harmonic bands. Specifically, we retain only the spectral components around integer multiples of the 360 kHz Qi fundamental frequency within a ±5 kHz bandwidth, estimate the background power from gesture-free recordings, and calculate the excess signal power caused by the hand.

Fig. 19 shows that the SNR decreases as the hand-to-charger distance increases, which is consistent with the near-field nature of the body-coupled WPT leakage signal. When the hand is close to the charger, the gesture-induced perturbation is clearly above the background level. As the distance increases, the coupling between the hand, body, and charging field becomes weaker, reducing the SNR margin. The measured curve remains above the 3 dB detection reference within approximately 20 cm, indicating that gesture-related perturbations are still

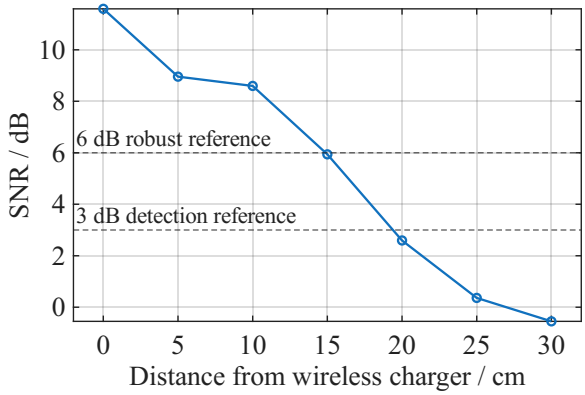


Fig. 19. Interaction-distance evaluation with a phone present on the wireless charging pad. The user's hand is placed 0–30 cm away from the charger at 5 cm intervals. SNR is computed over the selected 360 kHz-spaced harmonic bands after subtracting the background power. The signal remains detectable within approximately 20 cm, while interaction within 15 cm is recommended for more robust operation.

observable in this range. However, beyond 20 cm, the SNR approaches the background level, making the signal more vulnerable to environmental noise and vehicle-induced baseline fluctuations.

For robust in-vehicle interaction, we recommend using *MagHarp* within 15 cm of the wireless charger. Within this range, the SNR maintains a clearer margin above the detection reference and is closer to the robust operating region, providing more stable feature extraction for downstream gesture recognition. This range also matches natural center-console interaction behavior: the user's hand typically performs short mid-air gestures near the charging pad rather than far above it. Therefore, we regard 20 cm as the effective interaction range under our current setup, and recommend 15 cm as the practical operating distance for reliable use.

7.6 Phone-Absence Sensing

We further examine whether the wireless charger still exposes measurable signals when no phone is placed on the charging pad. This setting differs from normal charging because there is no active receiver, and the transmitter does not continuously emit a strong charging field. Instead, the charger periodically sends weak probing signals to detect whether a receiver is present. To observe this behavior, we collect time–frequency data when the user's hand is placed very close to the wireless charger.

Fig. 20 shows a representative time–frequency map in the phone-absence setting. We observe that the charger emits ping-like probing signals even without a phone. Each probing round contains two consecutive signal bursts, and the interval between two probing rounds is slightly less than 2 s. Although these probing signals are much weaker and more intermittent than the continuous leakage field during normal charging, they can still be received through the human-body coupling path when the user's hand is within approximately 5 cm of the charger.

These observations should be interpreted as a preliminary physical-channel validation rather than a mature phone-absence interaction function. The current ∼5 cm range only demonstrates that, even without an active phone receiver, the charger still exposes a measurable analog-ping channel that can be received through the human-body coupling path under very-close hand proximity. This result is therefore intended to verify the existence of the phone-absence sensing channel, not to claim that *MagHarp* already supports robust phone-absence gesture interaction at practical interaction distances. In the current prototype, phone-absence signals remain

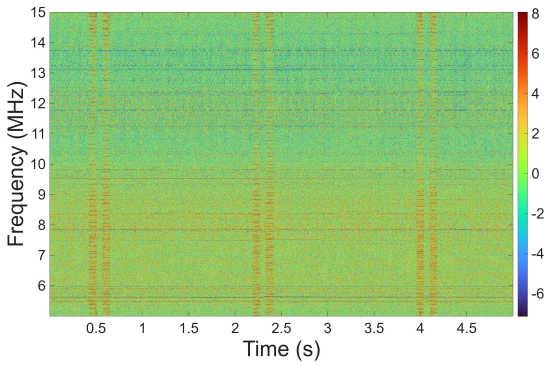


Fig. 20. Representative time–frequency map in the phone-absence setting. Without a phone on the charging pad, the wireless charger periodically emits ping-like probing signals. Each probing round contains two consecutive bursts, and adjacent probing rounds are separated by slightly less than 2 s. The signal is weak but can still be received through body coupling when the user's hand is within approximately 5 cm of the charger.

much weaker and more intermittent than the continuous leakage field observed during normal phone-present charging, which limits their usable range and temporal resolution.

Nevertheless, the structured nature of the ping signal suggests a promising direction for future work. Each probing round exhibits a repeatable burst pattern, which could be explicitly exploited by a dedicated ping-aware sensing pipeline. For example, future algorithms could synchronize feature extraction with individual analog-ping bursts, aggregate weak responses across consecutive probing rounds, and model the temporal evolution of hand-induced perturbations over multiple ping cycles. On the hardware side, the current wideband prototype could be replaced or augmented with a higher-gain low-noise analog front-end, narrower band-pass filtering around the ping-related spectral components, and improved shielding/grounding to increase the received SNR. These hardware and algorithmic improvements may extend the observable range beyond the current ∼5 cm preliminary result. We therefore position phone-absence sensing as an early feasibility finding and a future extension, while the main validated interaction mode of *MagHarp* remains the phone-present charging state evaluated in Sec. 7.5.

8 DISCUSSION AND LIMITATIONS

While *MagHarp* demonstrates robust performance in standard charging scenarios, the transition from a controlled prototype to a universally deployable automotive interface presents specific challenges regarding hardware heterogeneity, dynamic charging power states, and deployment miniaturization. Below, we discuss these challenges and outline technical pathways to address them.

8.1 Handling Hardware Heterogeneity via Cloud-Edge Synergy

Our current evaluation covers a representative range of devices; however, large-scale deployment faces the challenge of *WPT-Device Heterogeneity*. Although the Qi standard defines baseline protocols, varying manufacturer implementations (e.g., coil geometry, magnet shielding, and proprietary frequency tuning) lead to distinct electromagnetic leakage fingerprints. Furthermore, cross-ecosystem mismatches—such as charging an iPhone on a Xiaomi-proprietary pad—often result in negotiated power limits (e.g., reverting to 5W or 7.5W), which lowers the Signal-to-Noise Ratio (SNR) of the leakage field compared to matched high-power pairs. Currently, this variability often necessitates device-specific data collection. To overcome this and achieve *zero-shot* or *few-shot*

generalization, we envision the construction of a "WPT-Device Atlas" in the cloud. By aggregating technical specifications and signal profiles from public regulatory documents (e.g., FCC filings) and crowdsourced user data, we can build a lookup library of Charger-Vehicle-Phone combinations. When a user enters the car, the system identifies the phone model and charger ID. It then queries the cloud atlas to retrieve the expected fundamental frequency, harmonic distribution, and baseline power curve for that specific pair. This prior knowledge allows the local model to dynamically adjust its feature extraction filters and normalization parameters, minimizing or eliminating the need for user-specific calibration data.

8.2 Adaptive Sensing under Dynamic Power States

A unique characteristic of repurposing WPT for sensing is that the available carrier strength is coupled with the phone's charging state. Although the Qi2/MPP fundamental frequency remains relatively stable, the transmitted power can vary with state of charge (SoC), thermal throttling, phone alignment, and vendor-specific charging policies. As shown in our evaluation, *MagHarp* remains usable across low-, mid-, and high-SoC regimes, and the deployment-specific adaptation-profile mechanism can mitigate the performance drop in non-standard charging states. This indicates that charging-state variation mainly introduces a correctable baseline shift rather than destroying the gesture-induced perturbation. In the current prototype, we handle this variability through a stored calibration/adaptation profile and runtime background updates.

Future work can further reduce or even eliminate the need for labeled adaptation by incorporating phone-side charging context into the sensing pipeline. Modern smartphones can expose useful state information, such as battery level, charging status, temperature, charging mode, charging power, thermal state, and potentially Qi control messages through system APIs. Future versions of *MagHarp* could treat these variables as auxiliary inputs during training and jointly learn the relationship among gesture features, charging state, SoC, transmitted power, and thermal behavior. With sufficient training coverage over diverse charging conditions, the model could dynamically condition its normalization, feature weighting, filter selection, or classifier decision boundary on the current phone state. This would move *MagHarp* toward a zero-calibration deployment mode in which the system becomes usable immediately after the phone is placed on the wireless charger.

8.3 Deployment Path and Miniaturization Feasibility

We emphasize that the current *MagHarp* implementation is a proof-of-concept research prototype rather than a deployment-ready automotive module. While our present setup uses a HackRF One, a wideband amplifier, and a conductive-fabric pickup for flexibility and rapid signal exploration, the intended deployment path is substantially more compact. In our prototype, the signal acquisition electronics have already been consolidated into a compact enclosure that can be placed beneath the center console or underneath the seat, and the conductive electrode can be laminated into the seat upholstery or embedded beneath the seat surface. More importantly, a production realization does not require a full general-purpose SDR: *MagHarp* only needs to monitor a narrowband harmonic region around the selected carrier and extract low-rate amplitude/phase trajectories for gesture recognition. This functionality can be implemented with a dedicated narrowband analog front-end, a compact ADC, and an embedded microcontroller or automotive-grade RF SoC. With short shielded interconnects and local power regulation, the remaining electronics can be packaged as a small shielded module integrated into existing interior electronics. Therefore, we view *MagHarp* not as a literal zero-added-hardware system, but as a non-invasive, passive add-on sensing architecture that leaves the commodity charger and smartphone unchanged while remaining compatible with practical miniaturization.

9 CONCLUSION

As vehicles transform into sophisticated mobile computing environments, the tension between rich digital interaction and driving safety has reached a critical tipping point. In this paper, we presented *MagHarp* , a novel system that resolves this conflict not by modifying existing charging hardware, but by awakening the "blind" infrastructure already present in the center console: the wireless charging pad.

By reimagining the Wireless Power Transfer (WPT) field as an active sensing medium, *MagHarp* transforms a passive energy conduit into a unified, privacy-preserving gesture interface. Our work bridges the gap between electromagnetic physics and human-computer interaction; we derived a physics-informed impedance decoding framework that successfully isolates subtle, human-induced perturbations from the hostile, high-energy background of active charging. This approach overcomes the "Installation Paradox", enabling robust, eyes-free control of both vehicle infotainment and personal smartphones with lower additional sensing overhead than dedicated radar and without relying on always-on cameras.

Extensive evaluations across 5 commercial chargers, 4 smartphones, and longitudinal cross-day trials demonstrate that MagHarp achieves 97.6% average gesture recognition accuracy, with gesture-induced signals detectable within 20 cm and a recommended operating range within 15 cm. These results validate that the leakage fields surrounding us are not merely waste energy to be suppressed, but rich information channels waiting to be tapped. We hope this work inspires a new wave of Infrastructure-Opportunistic Sensing, paving the way for a future where ubiquitous interfaces are woven invisibly into the energy fabric of our daily lives — making interaction safer, touch-free, and truly worry-free.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (NSFC) under Grant No. 62502303, and by the Shanghai Municipal Natural Science Foundation under the 2024 “Science and Technology Innovation Action Plan” under Grant No. 24ZR1439100.

References

[1] American Automobile Association (AAA). 2024. Distracted Driving – AAA Exchange. [Online; accessed 2026-01-20]. https://exchange.aaa.com/safety/distracted-driving/

[2] Pew Research Center. 2015. The smartphone: An essential travel guide | Pew Research Center. [Online; accessed 2026-01-25]. https://www.pewresearch.org/short-reads/2015/04/14/smartphone-essential-travel-guide/

[3] Pew Research Center. 2015. The smartphone: An essential travel guide | Pew Research Center. [Online; accessed 2026-01-20]. https://www.pewresearch.org/short-reads/2015/04/14/smartphone-essential-travel-guide/

[4] Sang-Yoon Chang, Sristi Lakshmi Sravana Kumar, and Yih-Chun Hu. 2017. Cognitive wireless charger: Sensing-based real-time frequency control for near-field wireless charging. In *2017 IEEE 37th International Conference on Distributed Computing Systems (ICDCS)*. IEEE, 2302–2307.

[5] Zhaoxin Chang, Fusang Zhang, Jie Xiong, Weiyan Chen, and Daqing Zhang. 2024. MSense: Boosting wireless sensing capability under motion interference. In *Proceedings of the 30th Annual International Conference on Mobile Computing and Networking*. 108–123.

[6] Dongyao Chen, Mingke Wang, Chenxi He, Qing Luo, Yasha Iravantchi, Alanson Sample, Kang G Shin, and Xinbing Wang. 2021. MagX: wearable, untethered hands tracking with passive magnets. In *Proceedings of the 27th Annual International Conference on Mobile Computing and Networking*. 269–282.

[7] Tiantian Chen, Oscar Oviedo-Trespalacios, NN Sze, and Sikai Chen. 2022. Distractions by work-related activities: the impact of ride-hailing app and radio system on male taxi drivers. *Accident Analysis & Prevention* 178 (2022), 106849.

[8] Yihua Cheng, Yaning Zhu, Zongji Wang, Hongquan Hao, Yongwei Liu, Shiqing Cheng, Xi Wang, and Hyung Jin Chang. 2024. What do you see in vehicle? comprehensive vision solution for in-vehicle gaze estimation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 1556–1565.

[9] Gabe Cohn, Daniel Morris, Shwetak Patel, and Desney Tan. 2012. Humantenna: Using the Body as an Antenna for Real-Time Whole-Body Interaction. In *Proc. ACM CHI*. 1901–1910.

[10] Wireless Power Consortium. 2026. Home | Wireless Power Consortium. [Online; accessed 2026-01-27]. https://www.wirelesspowerconsortium.com/

[11] Team Counterpoint. 2025. Global In-Car Wireless Smartphone Charging: A Rapidly Expanding Market. [Online; accessed 2026-01-26]. https://counterpointresearch.com/en/insights/global-in-car-wireless-smartphone-charging-a-rapidly-expanding-market
[12] Minhao Cui, Binbin Xie, Qing Wang, and Jie Xiong. 2024. Evlesen: In-vehicle sensing with ev-leaked signal. In *Proceedings of the 30th Annual International Conference on Mobile Computing and Networking*. 679–693.
[13] Dian Ding, Lanqing Yang, Yi-Chao Chen, and Guangtao Xue. 2021. Vibwriter: Handwriting recognition system based on vibration signal. In *2021 18th Annual IEEE International Conference on Sensing, Communication, and Networking (SECON)*. IEEE, 1–9.
[14] Dian Ding, Lanqing Yang, Yi-Chao Chen, and Guangtao Xue. 2022. Handwriting recognition system leveraging vibration signal on smartphones. *IEEE Transactions on Mobile Computing* 22, 7 (2022), 3940–3951.
[15] Edison. 2024. The infinite dial. [Online; accessed 2026-01-25]. https://www.edisonresearch.com/wp-content/uploads/2024/06/Infinite-Dial-2024-Presentation.pdf
[16] Association for the Advancement of Automotive Medicine. 2024. Cellphone-Related Distracted Driving - Association for the Advancement of Automotive Medicine. [Online; accessed 2026-01-20]. https://www.aaam.org/education-resource-center/public-position-statements/cellphone-related-distracted-driving/
[17] Yongjian Fu, Lanqing Yang, Hao Pan, Yi-Chao Chen, Guangtao Xue, and Ju Ren. 2024. MagSpy: Revealing user privacy leakage via magnetometer on mobile devices. *IEEE Transactions on Mobile Computing* (2024).
[18] Guillermo Gallego, Tobi Delbruck, Garrick Orchard, Chiara Bartolozzi, Brian Taba, Gabriela Csurka, et al. 2022. Event-based vision: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 44, 1 (2022), 154–180.
[19] Danei Gong, Naiyu Zheng, Binbin Xie, Jie Xiong, Shuai Wang, Yuguang Fang, and Zhimeng Yin. 2025. SeRadar: Embracing Secondary Reflections for Human Sensing with mmWave Radar. In *Proceedings of the 31st Annual International Conference on Mobile Computing and Networking*. 231–246.
[20] Google Maps Help Community. 2024. How do I turn off the prompts to confirm road issues while I'm driving? - Google Maps Community. [Online; accessed 2026-01-20]. https://support.google.com/maps/thread/332360118/how-do-i-turn-off-the-prompts-to-confirm-road-issues-while-i-m-driving
[21] Daosong Hu, Mingyue Cui, and Kai Huang. 2025. FIFA: Fine-grained Inter-frame Attention for Driver's Video Gaze Estimation. In *Proceedings of the Computer Vision and Pattern Recognition Conference*. 18760–18769.
[22] Hua Huang, Hongkai Chen, and Shan Lin. 2019. Magtrack: Enabling safe driving monitoring with wearable magnetics. In *Proceedings of the 17th annual international conference on mobile systems, applications, and services*. 326–339.
[23] S. Y. Ron Hui and W. W. C. Ho. 2014. *Planar Wireless Charging Technology for Portable Electronic Products and Qi*. IEEE Press.
[24] Sakila S Jayaweera, Beibei Wang, and KJ Ray Liu. 2024. Robust In-Car Child Presence Detection using Commercial WiFi. In *Proceedings of the 30th Annual International Conference on Mobile Computing and Networking (MobiCom)*. 1799–1801.
[25] S Jeschke, M Maarleveld, J Baerenfaenger, B Schmuelling, and A Burkert. 2018. Challenges in EMC testing of EV and EVSE equipment for inductive charging. In *2018 International Symposium on Electromagnetic Compatibility (EMC EUROPE)*. IEEE, 967–971.
[26] Shenyao Jiang, Hao Zhou, Wangqiu Zhou, Xinyu Wang, Zhenjiang Li, and Yusheng Ji. 2025. FREAUTH^{+}: A Robust Frequency Feature-Based Device Authentication Mechanism for Magnetic Wireless Power Transfer System. *IEEE Transactions on Mobile Computing* (2025).
[27] Pascal Kowalczuk and Jennifer Musial. 2024. The impact of self-conscious emotions on the continuance intention of digital voice assistants in private and public contexts. *Computers in Human Behavior Reports* 15 (2024), 100450. doi:10.1016/j.chbr.2024.100450
[28] Akihiko Kumazawa, Yinliang Diao, Akimasa Hirata, and Hiroshi Hirayama. 2020. Reduction of Human Interaction with Wireless Power Transfer System Using Shielded Loop Coil. *Electronics* 9, 6 (2020). doi:10.3390/electronics9060953
[29] Ji Li, Yan Meng, Li Zhang, Guorui Chen, Ye Tian, and Haojin Xu. 2023. MagFingerprint: A Magnetic Based Device Fingerprinting in Wireless Charging. In *Proceedings of the IEEE INFOCOM 2023 - IEEE Conference on Computer Communications*. IEEE, 1–10.
[30] Y. Li, Y. Chen, L. Hu, K. Niu, and D. Zhang. 2023. M3Track: mmWave-based Multi-User 3D Posture Tracking with Uncertainty-Aware Motion Model. In *Proceedings of the 21st Annual International Conference on Mobile Systems, Applications, and Services (MobiSys)*. ACM, 1–13.
[31] Kang Ling, Haipeng Dai, Yunchao Liu, Alex X. Liu, Wei Wang, and Yu Gu. 2022. UltraGesture: Fine-Grained Gesture Sensing and Recognition. *IEEE Transactions on Mobile Computing* 21, 7 (2022), 2620–2636. Published online 2020.
[32] Haug Manuel, Beltle Michael, and Tenbohlen Stefan. 2021. Magnetic field emission of automotive inductive charging systems in the 9 kHz-30 MHz range. In *2021 Asia-Pacific International Symposium on Electromagnetic Compatibility (APEMC)*. IEEE, 1–4.
[33] Manuel Martin, Alina Roitberg, Monica Haurilet, Matthias Horne, Simon Reiß, Michael Voit, and Rainer Stiefelhagen. 2019. Drive&act: A multi-modal dataset for fine-grained driver behavior recognition in autonomous vehicles. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 2801–2810.
[34] Mayank. 2025. In car listening enters smartphone-first era: Futuresource research - RadioInfo Asia. [Online; accessed 2026-01-20]. https://radioinfo.asia/news/in-car-listening-enters-smartphone-first-era-futuresource-research/

[35] Pavlo Molchanov, Xiaodong Yang, Shalini Gupta, Kihwan Kim, Stephen Tyree, and Jan Kautz. 2016. Online detection and classification of dynamic hand gestures with recurrent 3d convolutional neural network. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 4207–4215.

[36] National Highway Traffic Safety Administration (NHTSA). 2023. Cell Phone Laws. Countermeasures That Work: A Highway Safety Countermeasure Guide for State Highway Safety Offices, 11th Edition. [Online; accessed 2026-01-20]. https://www.nhtsa.gov/book/countermeasures-that-work/distracted-driving/countermeasures/legislation-and-licensing/cell

[37] NCSL. 2026. Distracted Driving | Cellphone Use. [Online; accessed 2026-01-20]. https://www.ncsl.org/transportation/distracted-driving-cellphone-use

[38] Tao Ni, Xiaokuan Zhang, Chaoshun Zuo, Jianfeng Li, Zhenyu Yan, Wubing Wang, Weitao Xu, Xiapu Luo, and Qingchuan Zhao. 2023. Uncovering user interactions on smartphones via contactless wireless charging side channels. In *2023 IEEE Symposium on Security and Privacy (SP)*. IEEE, 3399–3415.

[39] Hao Pan, Feitong Tan, Wenhao Li, Yi-Chao Chen, Lanqing Yang, Guangtao Xue, and Xiaoyu Ji. 2022. MagDefender: Detecting eavesdropping on mobile devices using the built-in magnetometer. In *2022 19th Annual IEEE International Conference on Sensing, Communication, and Networking (SECON)*. IEEE, 28–36.

[40] Federal Register. 2013. Federal Register :: Visual-Manual NHTSA Driver Distraction Guidelines for In-Vehicle Electronic Devices. [Online; accessed 2026-01-20]. https://www.federalregister.gov/documents/2013/04/26/2013-09883/visual-manual-nhtsa-driver-distraction-guidelines-for-in-vehicle-electronic-devices

[41] Status Report. 2019. Untitled. [Online; accessed 2026-01-25]. https://www.iihs.org/api/datastoredocument/status-report/pdf/54/1

[42] ResearchAndMarkets.com. 2024. China Passenger Car Mobile Phone Wireless Charging Research Report 2024: Over 80% of Automotive Mobile Phone Wireless Charging Modules have Obtained Qi 2.0 Certification. Business Wire. Accessed: 2026-01-27. https://www.businesswire.com/news/home/20241129404402/en/China-Passenger-Car-Mobile-Phone-Wireless-Charging-Research-Report-2024-Over-80-of-Automotive-Mobile-Phone-Wireless-Charging-Modules-have-Obtained-Qi-2.0-Certification---ResearchAndMarkets.com

[43] Munehiko Sato, Ivan Poupyrev, and Chris Harrison. 2012. Touché: Enhancing Touch Interaction on Humans, Screens, Liquids, and Everyday Objects. In *Proc. ACM CHI*. 483–492.

[44] Haley M So, Laurie Bose, Piotr Dudek, and Gordon Wetzstein. 2024. Pixelrnn: in-pixel recurrent neural networks for end-to-end-optimized perception with neural sensors. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 25233–25244.

[45] Yuqi Su, Fusang Zhang, Beihong Jin, and Daqing Zhang. 2025. Manipulation of Acoustic Focusing for Multi-target Sensing with Distributed Microphones in Smart Car Cabin. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 9, 2 (2025), 1–28.

[46] Yuqi Su, Fusang Zhang, Kai Niu, Tianben Wang, Beihong Jin, Zhi Wang, Yalan Jiang, Daqing Zhang, Lili Qiu, and Jie Xiong. 2024. Embracing distributed acoustic sensing in car cabin for children presence detection. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 8, 1 (2024), 1–28.

[47] S. Sun, A. P. Petropulu, and H. V. Poor. 2020. MIMO Radar for Advanced Driver-Assistance Systems and Autonomous Driving: Advantages and Challenges. *IEEE Signal Processing Magazine* 37, 4 (2020), 101–117. doi:10.1109/MSP.2020.2993208

[48] Uber. 2024. Getting a trip request | Driving & Delivering | Uber Help. [Online; accessed 2026-01-20]. https://help.uber.com/driving-and-delivering/article/getting-a-trip-request?nodeId=7162bba3-ba0b-4f76-b900-645bf43c7b07

[49] Edward Jay Wang, Jake Garrison, Eric Whitmire, Mayank Goel, and Shwetak Patel. 2017. Carpacio: Repurposing capacitive sensors to distinguish driver and passenger touches on in-vehicle screens. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology*. 49–55.

[50] Shuai Wang, Yunze Zeng, Vivek Jain, and Parth Pathak. 2025. UMusic: In-car Occupancy Sensing via High-resolution UWB Power Delay Profile. In *Proceedings of the 23rd ACM Conference on Embedded Networked Sensor Systems*. 116–129.

[51] Wei Wang, Alex X Liu, Muhammad Shahzad, Kang Ling, and Sanglu Lu. 2017. Device-Free Human Activity Recognition Using Commercial WiFi Devices. *IEEE J. Sel. Areas Commun.* 35, 5 (2017), 1118–1131.

[52] Chongzhi Xu, Xiaolong Zheng, Zhiyuan Ren, Liang Liu, and Huadong Ma. 2024. Uhead: Driver attention monitoring system using uwb radar. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 8, 1 (2024), 1–28.

[53] Dingkang Yang, Shuai Huang, Zhi Xu, Zhenpeng Li, Shunli Wang, Mingcheng Li, Yuzheng Wang, Yang Liu, Kun Yang, Zhaoyu Chen, et al. 2023. Aide: A vision-driven multi-view, multi-modal, multi-tasking dataset for assistive driving perception. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 20459–20470.

[54] Lanqing Yang, Yi-Chao Chen, Hao Pan, Dian Ding, Guangtao Xue, Linghe Kong, Jiadi Yu, and Minglu Li. 2020. Magprint: Deep learning based user fingerprinting using electromagnetic signals. In *IEEE INFOCOM 2020-IEEE Conference on Computer Communications*. IEEE, 696–705.

[55] Zihao Zhan, Yeton Yang, Hui Shan, Hanyu Wang, and Yier Jin. 2024. VoltSchemer: Use Voltage Noise to Manipulate Your Wireless Charger. In *Proceedings of the 33rd USENIX Security Symposium (USENIX Security)*. USENIX Association.

[56] Fusang Zhang, Zhaoxin Chang, Jie Xiong, Junqi Ma, Jiazhi Ni, Beihong Jin, and Daqing Zhang. 2022. Embracing consumer-level UWB-equipped devices for fine-grained wireless sensing. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 6, 4 (2022), 1–27.

[57] Yi Zhang, Weiying Hou, Zheng Yang, and Chenshu Wu. 2023. {VeCare}: Statistical acoustic sensing for automotive {In-Cabin} monitoring. In *20th USENIX Symposium on Networked Systems Design and Implementation (NSDI 23)*. 1185–1200.

[58] Juncen Zhu, Jiannong Cao, Yanni Yang, Wei Ren, and Huizi Han. 2023. mmDrive: Fine-Grained Fatigue Driving Detection Using mmWave Radar. *ACM Transactions on Internet of Things* 4, 2 (2023), 1–26.