RadSee: See Your Handwriting Through Walls Using FMCW Radar

Shichen Zhang, Qijun Wang, Maolin Gan, Zhichao Cao, and Huacheng Zeng
Department of Computer Science and Engineering, Michigan State University
{sczhang, qjwang, ganmaoli, caozc, hzeng}@msu.edu

***Abstract*—This paper aims to design and implement a radio device capable of detecting a person's handwriting through a wall. Although there is extensive research on radio frequency (RF) based human activity recognition, this task is particularly challenging due to the *through-wall* requirement and the *tiny-scale* handwriting movements. To address these challenges, we present RadSee—a 6 GHz frequency modulated continuous wave (FMCW) radar system designed for detecting handwriting content behind a wall. RadSee is realized through a joint hardware and software design. On the hardware side, RadSee features a 6 GHz FMCW radar device equipped with two custom-designed, high-gain patch antennas. These two antennas provide a sufficient link power budget, allowing RadSee to "see" through most walls with a small transmission power. On the software side, RadSee extracts effective phase features corresponding to the writer's hand movements and employs a bidirectional LSTM (BiLSTM) model with an attention mechanism to classify handwriting letters. As a result, RadSee can detect millimeter-level handwriting movements and recognize most letters based on their unique phase patterns. Additionally, it is resilient to interference from other moving objects and in-band radio devices. We have built a prototype of RadSee and evaluated its performance in various scenarios. Extensive experimental results demonstrate that RadSee achieves 75% letter recognition accuracy when victims write 62 random letters, and 87% word recognition accuracy when they write articles.**

I. Introduction

Handwriting on physical papers and electronic devices (e.g., iPad) is one of the most common human activities to keep and transfer information. Even in today's digital world, people tend to write more than they might think [3]. In some scenarios, the confidentiality of written content is of paramount importance. A natural question to ask is: *if one is writing important documents on a desk in a private room, is it possible for an attacker outside the room to detect the letters being written through the wall?* Understanding the capability and performance limits of such an attacker would inform the public of not only potential threats but also possible countermeasures, thereby preventing information leakage and enhancing human activity privacy.

Recent years have witnessed significant progress in remote human activity recognition (HAR) using different sensing technologies such as cameras [11], [22], [33], ultrasound [36], [56], Wi-Fi [13], [21], [27], [28], [38], [45], RFID [51], [59], and millimeter-wave (mmWave) radar [4], [19], [31], [32], [43], [53], [57], [62]. In contrast to existing work, the task of detecting handwriting content through a wall is unique yet challenging in the following three aspects.

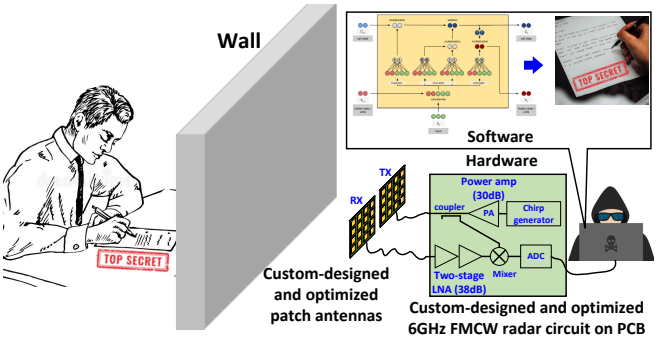


Fig. 1: RadSee is a joint hardware (radar) and software (deep learning) design to detect the letters being written by a victim behind a wall.

(a) Through-wall detection. This requirement significantly limits the viable sensing techniques for this task. Camera-based computer vision (CV) can be used for human activity recognition by analyzing video data to identify and classify different human actions and movements [22], [33]. Powered by advanced deep neural network (DNN) techniques, a camera system can easily recognize the handwriting characters from a distance [14], [40]. However, camera-based HAR systems are limited by occlusions and thus not applicable to through-wall detection. Ultrasound sensors have also been widely used for HAR. They emit high-frequency sound waves that bounce off objects and produce echoes, which can be analyzed to determine the patterns of human activities [12], [36], [56]. Unlike camera sensors, ultrasound can work even in low light conditions and does not require a line-of-sight path. But ultrasound has a very limited ability of passing through walls due to its short wavelength, making it unsuitable for this task. Radio frequency (RF) has emerged as a popular technology for HAR such as gesture recognition [13], [16], [23], [27], [38], [47], [61], keystroke detection [5], [58], and vital signal detection [15], [48]. Among existing RF technologies, high-frequency signals (e.g., mmWave) have very limited ability to pass through a wall. Therefore, radio signals on sub-10 GHz bands appear to be the plausible carrier for HAR behind walls.

(b) Millimeter-level hand movement for writing. Hand-

Network and Distributed System Security (NDSS) Symposium 2025
24-28 February 2025, San Diego, CA, USA
ISBN 979-8-9894372-8-3
https://dx.doi.org/10.14722/ndss.2025.230322
www.ndss-symposium.org

writing features very small movements compared to other human activities. Typically, the movement of a pen-holding hand is smaller than 1 cm for both paper and iPad writings. When using an RF system for handwriting detection, its detection resolution is determined by its signal wavelength. On one hand, high-frequency mmWave (e.g., 60 GHz and 77 GHz) signals are capable of detecting sub-mm movement of an object but cannot pass through a wall. On the other hand, low-frequency (e.g., 915 MHz) microwave signals can easily pass through a wall but cannot detect the mm-level movement of an object. On the middle-frequency spectrum bands such as 2.4 GHz and 5 GHz, channel state information (CSI) in Wi-Fi networks has been extensively used for HAR [13], [21], [27], [28], [45], [50]; and its application on 5 GHz frequency bands seems a possible solution to achieve the desired trade-off between wall penetration and detection resolution. However, Wi-Fi CSI-based HAR is a *non-coherent* detection approach that suffers from phase, frequency and timing misalignments in hardware. As such, it is incapable of detecting mm-level movement in time. Recently, 6 GHz FMCW radar, which is a coherent detection system, has been used for HAR such as human body skeleton construction [1], [29], [63]–[65]. This approach uses custom-designed hardware and promises high accuracy and stability. However, so far, its applications are limited to the detection of large-scale movements such as people walking and interaction.

(c) Interference resilience. The detection of handwriting may suffer from interference from other moving objects such as a walking person around the writer. It may also suffer from interference from indirect paths between the writer and the detection equipment. Actually, such interference is a notorious issue with RF sensing [25], [37], [55]. This issue is particularly acute in sub-6 GHz RF sensing systems. If not addressed, the interference may appear dominant and place a fundamental limit on the detection performance. Moreover, since different scenarios have different multi-path effects and different moving objects/people, addressing the interference is critical to extract environment-independent features for handwriting recognition and ultimately develop a radio detector that can work in new environments.

In this paper, we present RadSee, a 6 GHz FMCW radar system for detecting the handwriting activities behind a wall, as shown in Fig. 1. RadSee is realized through a joint hardware and software design. In terms of hardware, RadSee builds a 6 GHz FMCW radar with highly optimized patch antennas. In terms of software, RadSee first extracts the phase information of demodulated FMCW signals and employs a deep neural network (DNN) model for letter classification. Combining the hardware and software innovations, RadSee is capable of continuously detecting mm-level handwriting movement over time and recognizing most letters based on their unique phase patterns.

RadSee addresses **Challenge (a)** with its FMCW modulation, its high-gain patch antenna, and its optimized baseband analog filter. RadSee has co-located Tx and Rx RF chains, making it possible to perform coherent signal demodulation for handwriting recognition. In addition, the optimized patch antennas have a total 36 dBi gain for wall penetration. RadSee addresses **Challenge (b)** by using the *phase* information of demodulated FMCW signals to extract the features of handwriting movements. FMCW radar has been widely used for ranging. Its range resolution is $\frac{c}{2B}$, where c is light speed and B is signal bandwidth. Achieving the range resolution of 1 mm requires $B = \frac{3\times10^8}{2\times10^{-3}} = 150$ GHz signal bandwidth, which is impossible in practice. However, the *phase* of demodulated FMCW signals is much more sensitive to the movement of an object. In theory, 1 mm hand movement corresponds to 14° phase change of the demodulated signal, which is easy to detect. Therefore, RadSee uses the *phase* of demodulated FMCW signals as the features of letter recognition. RadSee addresses **Challenge (c)** by demodulating the reflective signals only from the handwriting movement. This is achieved by its FMCW modulation and high-directional patch antennas. The FMCW modulation allows it to focus on the Range-FFT bin that corresponds to the distance of interest; the patch antennas allow it to focus on the reflective signal from the direction of interest. Combining FMCW modulation and antenna directivity, RadSee is capable of detecting a clear phase pattern corresponding to the handwriting movements behind a wall using a small transmission power (20 dBm).

Based on the demodulated FMCW signals, RadSee employs a bidirectional long short-term memory (BiLSTM) model to classify the handwriting characters (a-z, A-Z, and 0-9). Different from other human activities such as keystroke [5], [58], handwriting is a smooth and continuous movement of the pen-holding hand. As such, handwriting tends to generate a unique temporal phase pattern for each letter. That is the reason why RadSee uses BiLSTM to classify a phase data sequence. Of a phase data sequence, some parts may be very important for letter recognition (e.g., those turning points), while some parts may not carry useful information (e.g., horizontal strokes). Therefore, RadSee adds an attention layer to the BiLSTM model so that the model can automatically focus on those important parts of a phase data sequence for letter classification. Powered by the BiLSTM model and its attention mechanism, RadSee is capable of recognizing handwriting letters based on their unique movement patterns.

We have built a prototype of RadSee (through PCB fabrication) and evaluated its performance in various scenarios. Experimental results show that, when placed behind office interior drywalls and external wood/vinyl walls, RadSee achieves 75% letter recognition accuracy when victims randomly write 62 different letters and 87% word recognition accuracy when victims write articles. Notably, RadSee demonstrates its resilience to the interference from walking persons around the victim writer and the interference from other radio devices.

Table I shows the comparison of RadSee and its related work. It advances the state-of-the-art in the following aspects.

- It designs and implements a 6 GHz FMCW radar device that can detect mm-level movements of an object behind a wall using a small transmission power.

TABLE I: Related work on human activity recognition. $\mathcal{W}$ = "See through wall?", $\mathcal{M}$ = "Mm-level movement detection?", $\mathcal{R}$ = "Resilient to multipath?", $\mathcal{I}$ = "Resilient to interference from other moving objects?", $\mathcal{S}$ = "Classification size".

References	Objective	Technique	$\mathcal{W}$	$\mathcal{M}$	$\mathcal{R}$	$\mathcal{I}$	$\mathcal{S}$
RF-Capture [1], RF-Avatar [64], RF-Pose [63], RF-Pose3D [65], RF-Action [29]	Human body skeleton	6GHz FMCW radar	✓	✗	✓	✓	N/A
WiSIA [28], WiPose [21], F. Wang [45]	Radio imaging	Wi-Fi	✗	✗	✗	✗	N/A
Tadar [59], RF-HMS [51]	Human tracking	RFID	✓	✗	✗	✗	N/A
mtrack [53]	Hand writing	mmWave	✗	✓	✗	✓	N/A
WiKey [5]	Key stroke	Wi-Fi	✗	✗	✗	✗	37
WiHF [27], WriFi [13], WiSee [38]	Gesture recognition	Wi-Fi	✗	✗	✗	✗	26
Soli [32]		mmWave	✗	✓	✗	✓	4
PhaseBeat [48]	Vital sign	Wi-Fi	✓	✗	✗	✗	N/A
RF-SCG [15]		mmWave	✗	✓	✗	✓	N/A
RadSee (ours)	**Hand writing**	**FMCW radar**	✓	✓	✓	✓	**62**

- RadSee is capable of detecting the letters that one is writing behind a wall. Furthermore, it is resilient to the interference from other mobile objects and other radio devices.
- Extensive experimental results show that RadSee can achieve over 75% accuracy when detecting 62 random letters and 87% word recognition accuracy behind walls.

II. Attack Model

Attack Scenarios. We consider a scenario as shown in Fig. 1, where one is writing a confidential document on a paper or an electronic device (e.g., iPad and Kindle Scribe) in a private room (e.g., government office, business office, hotel room, and apartment). Inside the room there may be other static objects (e.g., furniture) and people performing daily activities. Outside the room there is a malicious attacker who aims to detect the content (English letters and Arabic numbers) being written by the victim.

Attacker's Assumptions. We assume that the attacker has physical access to the space behind the wall which the victim is facing toward. As such, the radio signals for detecting the victim's handwriting movements would not be blocked by the victim's torso. We also assume that the attacker knows the layout of the room and the approximate location of the victim. However, the accurate location of victim's writing hand will be estimated by the attacker using radar signal. We know that it is not improbable to obtain the knowledge about the location of a desk in a room, as many public spaces such as hotels have standard layouts that are consistent across rooms. Furthermore, we assume that there are no RF-shielding materials inside the wall between victim and the attacker.

Challenges. As we stated before, there are three grand challenges that must be addressed for the design of such an adversarial device, namely, through-wall detection, mm-level recognition, and interference resilience. In addition, handwriting recognition has 62 character candidates (26 low-case letters, 26 upper-case letters, and 10 Arabic numbers) for classification. Such a large character set adds another level of challenge to the task. To address these challenges, it calls for a joint hardware and software design for such an attack device.

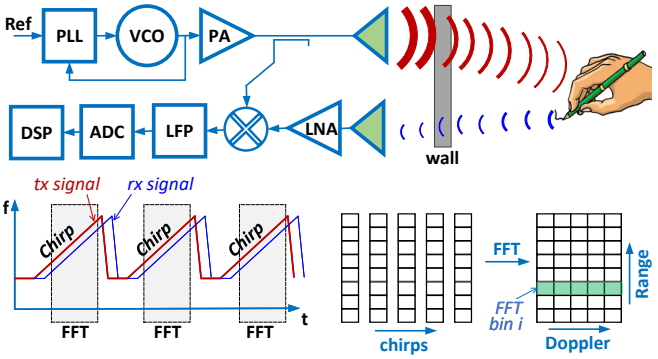


Fig. 2: Illustration of radar device.

III. RadSee: Design Analysis

A. A Primer on FMCW Radar

FMCW radar is an active radio device that uses frequency modulation to generate a continuous wave signal with a linear frequency sweep. This signal is transmitted from the radar antenna toward a target, and the reflection from the target is received by the radar antenna. The frequency difference between the transmitted and received signals, known as the beat frequency, is proportional to the range of the target. By analyzing the beat frequency over time, FMCW radar can determine the distance and velocity of the target.

Fig. 2 shows the diagram of an FMCW radar device. It transmits frequency-modulated continuous-wave signals and receives the reflective signals from the surrounding objects. Denote $s_T(t)$ as the transmitting signal and $s_R(t)$ as the received echo from an object. Mathematically, we have

$$s_T(t) = \exp\left(j(2\pi f_0 t + \pi K t^2)\right), \tag{1}$$

and

$$s_R(t) = \alpha s_T(t - 2d/c), \tag{2}$$

where f_0 is the starting frequency, K is the frequency ramp rate, α is the path attenuation, d is the distance from the radar to the object of interest, and c is light speed.

The received signal and the transmitted signal are mixed together, generating the intermediate frequency (IF) signal. The IF signal can be written as:

$$s_{IF}(t) = s_T(t)s_R(t)^* = \exp\Big(\underbrace{j4\pi K\frac{d}{c}t}_{\text{frequency}} + \underbrace{j4\pi f_0\frac{d}{c}}_{\text{phase}} - \underbrace{j4\pi K\frac{d^2}{c^2}}_{\text{negligible}}\Big). \tag{3}$$

As we can see from (3), the observed frequency and phase both contain the distance information. Typically, the frequency

term in (3) is used to estimate the range of an object, while the phase term is used to estimate the velocity of the object. Specifically, the range and velocity of the object are estimated as follows.

- **Range.** As illustrated in Fig. 2, the IF signal from each chirp is digitized and converted to the frequency domain through FFT operation. Suppose that the FFT size is N and a peak is identified at the ith FFT bin ($0 \leq i \leq N-1$). Then, the distance of the corresponding object is $d = \frac{c}{2B} i$, where B is the FMCW signal bandwidth. Accordingly, the range resolution is $\Delta d = \frac{c}{2B}$, which is determined solely by the FMCW signal bandwidth.
- **Velocity.** Grouping an array of chirps together, the velocity of the object can be accurately estimated by performing the second FFT operation on the ith Range-FFT bins. Suppose that the time duration of one chirp is T and that the FFT size is M. In this case, a peak is identified at the kth FFT bin, which allows us to calculate the velocity of the object $v = \frac{kc}{2MTf_0}$. Accordingly, the velocity resolution is $\Delta v = \frac{c}{2MTf_0}$, which is determined by three parameters: the initial frequency, the time duration of a chirp, and the number of used chirps.

B. Feasibility Analysis

To detect fine-grained movements, the first option that came to our mind is mmWave FMCW radar, which is widely available on market at a low price. Particularly, existing work (e.g., [19], [31], [43]) has demonstrated the ability of mmWave radars to "see" through walls made of cotton and glass. A key question to ask is whether a mmWave radar can "see" through typical walls in our daily lives. To answer this question, we conducted experiments using IWR1642BOOST 77 GHz mmWave FMCW radar from Texas Instruments (TI) with a bandwidth of 1.1 GHz. We placed the mmWave radar behind an office drywall to detect the handwriting in a room. Fig. 3 shows our writing content. Fig. 4(a) shows experimental setting and the corresponding FFT-bin's amplitude and phase over time. We did not observe any amplitude or phase changes over time caused by the handwriting. This indicates that mmWave signals cannot go through the drywall under test.

Another possible approach to this task is to use Wi-Fi-based channel state information (CSI). Since Wi-Fi uses 2.4 GHz and 5 GHz frequency bands, its signal is able to penetrate walls for movement detection. To examine this approach, we conducted experiments in the same scenario as the previous case. Fig. 4(b) shows the measured CSI at a receiver when using Wi-Fi channel #3 (2412 MHz–2432 MHz). We observed random CSI changes over time, and did not find any patterns on the CSI's amplitude and phase that are related to the handwriting movement. Similar results are observed for the CSI measured on Wi-Fi channel #36 (5170 MHz–5190 MHz). This can be attributed to the non-coherent detection of a Wi-Fi receiver. Since Wi-Fi transmitter and receiver are driven by different clocks, the measured CSI suffers from carrier frequency and sampling time offsets, making it unreliable to extract the pattern of tiny-scale movements.

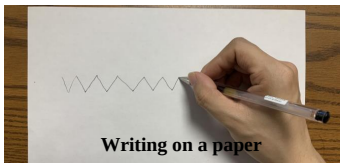


Fig. 3: Illustration of handwriting pattern.

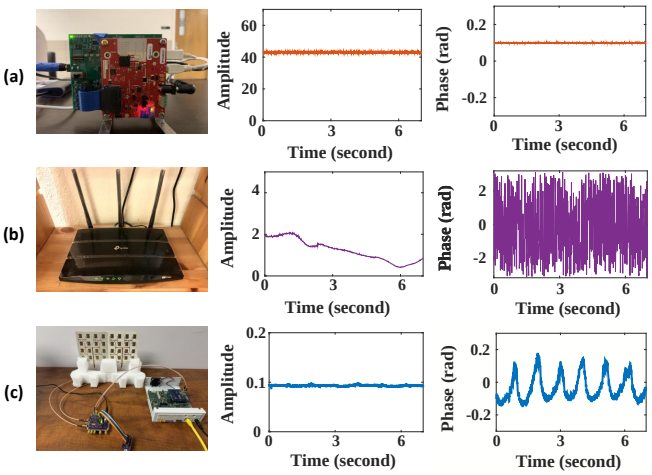


Fig. 4: (a) The amplitude and phase of the corresponding FFT-bin from IWR1642BOOST mmWave FMCW radar. (b) The amplitude and phase of a subcarrier from a Wi-Fi receiver. (c) The amplitude and phase of the corresponding FFT-bin from RadSee.

In comparison, we replaced the mmWave/Wi-Fi device with RadSee—our custom-designed 6 GHz FMCW radar. Fig. 4(c) shows the corresponding FFT-bin's amplitude and phase over time. It can be seen that the phase pattern is significant and that the phase pattern is consistent with the handwriting trajectory on the paper (see Fig. 3). This demonstrates the ability of RadSee to "see" through the wall under test.

Why Use 6 GHz FMCW Radar? Some may inquire about the suitability of other frequencies for through-wall and fine-grained movement detection. Low-frequency (0–3 GHz) radio signals have large wavelengths, rendering them incapable of detecting movements at the millimeter level. High-frequency (20–300 GHz) radio signals, on the other hand, have a large path loss and a significant penetration loss; thus they cannot travel through walls with normal transmission power. Radio signals in the range from 3 GHz to 20 GHz, however, should be suitable for this task. We opted for 6 GHz due to the availability and cost-effectiveness of electronic chips for FMCW radar implementation, including phase-locked loop (PLL), voltage-controlled oscillator (VCO), power amplifier (PA), low-noise amplifier (LNA), etc. On the market, only 6 GHz chips are available for individual customers at a reasonable price, thanks to the widespread production of 5 GHz Wi-Fi industry. The cost of our prototype is approximately $500.

Millimeter-level Movement Detection. If an FMCW radar wants to achieve 1 mm range resolution, it will need 150 GHz spectrum bandwidth, which is impossible in practice. There-

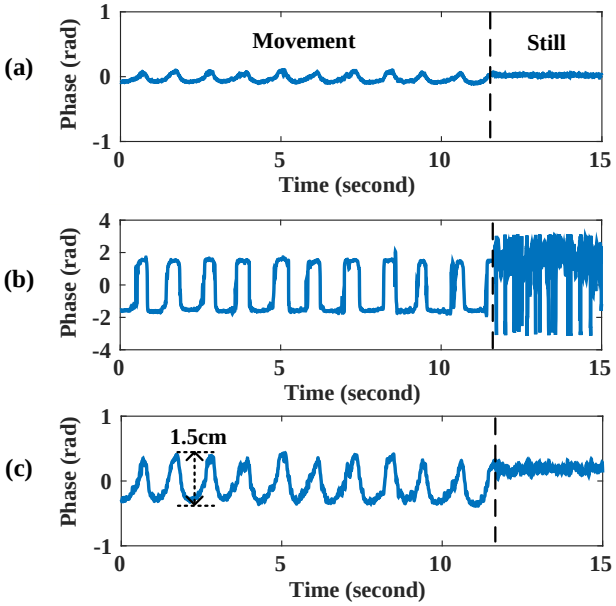


Fig. 5: Phase observations at a behind-wall radar when one is continuously writing back-and-forth on a paper within 1.5 cm. (a) phase data before DC component removal. (b) phase data after DC component removal. (c) phase data after partial DC component removal.

fore, RadSee uses the phase information of its demodulated FMCW signal to infer the movement pattern of handwriting. Based on Eqn. (3), when the object moves 1 mm, RadSee will observe $\frac{2df_0}{c}2\pi = 0.25$ radian (about $14°$) phase change on the corresponding Range-FFT bin. Typically, handwriting movement is larger than 5 mm, which will generate $70°$ phase change on the Range-FFT bin. Therefore, the radar will measure the phase pattern over time when a victim is writing, and use the temporal phase pattern to classify the letters being written.

Fig. 5 shows the observed phase change of a Range-FFT bin when one is writing back and forth on a paper behind a thick office drywall. The distance between the writing hand and the wall is about 2 m. The radar was placed on the other side of the wall, with a distance of 0.5 m. The person wrote back and forward within a vertical range of 1.5 cm. It can be observed from Fig. 5(a) that the phase changes as the pen-holding hand moves. However, the phase dynamic range is small. The small dynamic range is attributed to a DC voltage component of the received signal. The DC component, which can be modeled as a constant complex number, is the reflective signals from static objects (e.g., furniture and human body) of the same distance. Fortunately, the DC component is static over time and thus can be easily removed. Ideally, we should completely remove the DC component to maximize the phase sensitivity. However, when we completely remove the DC component, the time period of no-movement will have an irregular phase pattern as illustrated in Fig. 5(b), making it hard for RadSee to detect the gap between two consecutive letters. Therefore,

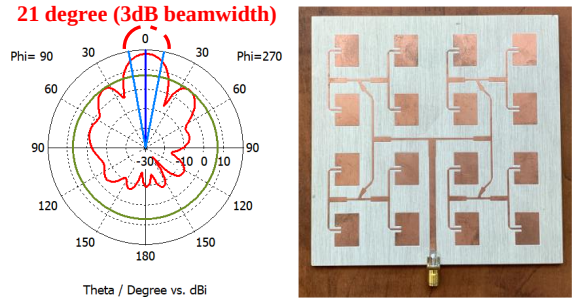


Fig. 6: The gain pattern of the patch antenna (left). The custom-designed directional antenna (right).

we partially remove the DC component to strike a balance between movement detection sensitivity and the phase stability of non-movement periods. Fig. 5(c) shows the observed phase of a Range-FFT bin over time.

Interference Resilience to Other Mobile Objects. In the proximity of a target writer, there may be many static objects such as desks, chairs, books, and lamps. Fortunately, the static objects will not generate interference for the detection of RadSee as their reflective signals appear to be a constant complex number (DC component) over time. Such a constant can be easily removed or adjusted to extract the useful phase information. As stated before, RadSee may suffer from interference from two sources: *(i) channel multi-path*, and *(ii) movement of other objects* (e.g., a walking person). Actually, RadSee is resilient to the interference from these two sources, thanks to its FMCW modulation and antenna directivity. We explain the reasons below.

- *FMCW Modulation (Distance Filter).* If two moving objects have different distances to the radar and their range difference is larger than the radar's range resolution, their phase-change patterns will appear on different Range-FFT bins and will not interfere with each other. Therefore, increasing the range resolution of RadSee is critical for reducing the interference from mobile objects. RadSee uses 1.1 GHz (5.4-6.5 GHz) bandwidth and thus has a range resolution of 14 cm. This means that, if separated by 14 cm, a mobile object (e.g., writer's chest movement of breathing) will not generate interference to RadSee's handwriting detection.
- *Patch-array Antenna (Directional Filter).* In addition to offering high link gain, the patch-array antenna also serves as a directional filter to suppress the interference from undesired azimuth/elevation angles. We designed and optimized the patch-array antenna using CST Studio Suite [9] and fabricated the path-array antenna as shown in Fig. 6. The main lobe of the antenna has an angular width of $21°$ (3 dB), which means that this antenna can effectively mitigate interference from mobile objects when they are positioned $21°$ or more away from the writer.

Combining its FMCW modulation and patch-array antenna, RadSee is capable of extracting the phase information corresponding to the movement within a small spot of interest,

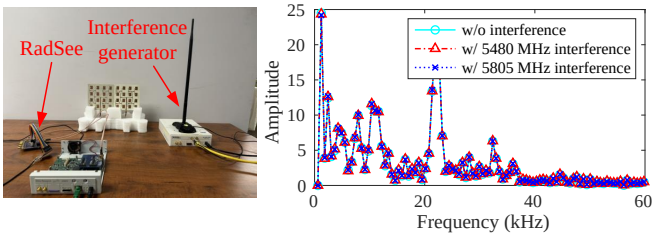


Fig. 7: Study an FMCW radar's resilience to radio interference from Wi-Fi devices: experimental setup (left) and experimental results (right).

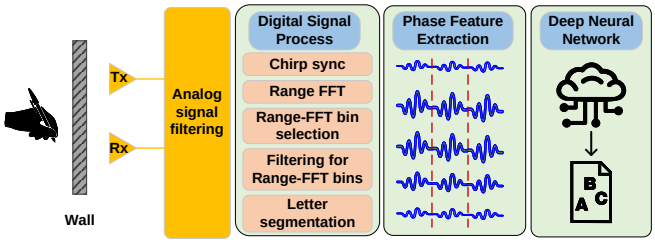


Fig. 8: RadSee process overview.

while being resilient to interference from other moving objects.

Interference Resilience to In-band Wi-Fi Devices. Although RadSee operates on a frequency band that overlaps with 5 GHz Wi-Fi, it differs significantly from Wi-Fi in two key aspects. First, RadSee has a bandwidth of 1.1 GHz, while Wi-Fi devices typically operate within a bandwidth of 20 or 40 MHz. Second, RadSee utilizes an FMCW waveform, whereas Wi-Fi devices use an Orthogonal Frequency-Division Multiplexing (OFDM) waveform. OFDM waveforms are characterized by pseudo-noise-like signals. When an OFDM signal is correlated with an FMCW signal over time, the correlation result is nearly zero. Therefore, in theory, RadSee is resilient to radio interference from the Wi-Fi devices in its proximity.

To validate the above theory, we conducted experiments by observing RadSee's IF signals in two cases: with and without radio interference from a Wi-Fi device, as shown in Fig. 7. To better control the experiments, we use a Universal Software Radio Peripheral (USRP) device for continuous Wi-Fi signal generation at two frequencies: 5.480 GHz and 5.805 GHz. The bandwidth of Wi-Fi signals is 20 MHz. The scene is static during the experiments. Fig. 7 presents RadSee's IF signals (i.e., the input of DNN) in three cases: i) no radio interference from the Wi-Fi device, ii) radio interference from 5.480 GHz Wi-Fi device, and iii) radio interference from 5.805 GHz Wi-Fi device. It can be seen that the IF signals generated by RadSee are almost the same in these three cases. This indicates that RadSee is resilient to radio interference from Wi-Fi devices.

IV. RadSee: Data Processing

In this section, we present the signal processing pipeline of RadSee, as outlined in Fig. 8. We first elaborate on the signal processing modules for phase feature extraction and then use k-nearest neighbor (kNN) to validate the extracted features.

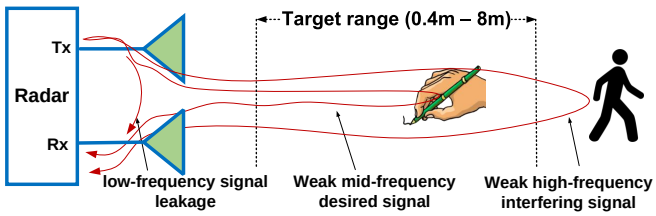


Fig. 9: Illustration of the received signals at the radar.

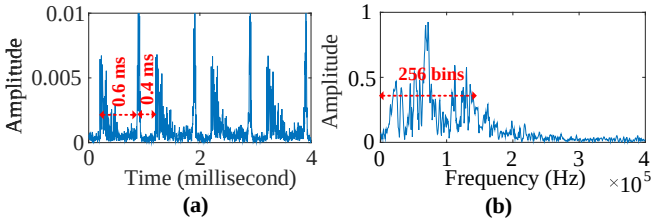


Fig. 10: Illustration of the IF signal. (a) the IF signal in time domain. (b) the IF signal after FFT operation.

A. Signal Processing

Analog Signal Filtering. The received signal at RadSee may have different components, including RF leakage on PCB, desired echo from handwriting, and undesired echo from other moving objects, as shown in Fig. 9. Since the RF leakage signal is very close to zero frequency, RadSee uses a high-pass filter with 5 kHz cutoff frequency to suppress the RF signal leakage. Meanwhile, the undesired high-frequency signal from other moving objects may generate interference to the desired signal if not suppressed in the analog domain. To do so, RadSee employs a first-order low-pass filter with a bandwidth of 100 kHz for the suppression of high-frequency echoes from undesired moving objects. Combining the high-pass and low-pass filters, RadSee has a band-pass filter from 5 kHz to 100 kHz, corresponding to a target range from 0.4 m to 8 m for handwriting detection.

Range-FFT. RadSee sets its chirp cycle time to 1 ms. For each chirp cycle, RadSee sets its transmission time to 0.6 ms and idle/delay time to 0.4 ms as shown in Fig. 10(a). As the PLL and VCO are typically not very stable at the beginning and end of their frequency ramping, RadSee discards 0.05 ms at the beginning and at the end of its transmission period, resulting in only 0.5 ms for useful signal reception. To best observe this useful signal in the digital domain, RadSee samples its received signal at 5 MSps. As a result, it obtains 2,500 complex samples from each chirp cycle. To further improve the range resolution, RadSee adds zeros behind the 2,500 samples to perform 8,192-point Range-FFT operation. The resultant Range-FFT bins are shown in Fig. 10(b). Of the resulting 8,192 bins, only the first 256 are under examination.

Filtering for Range-FFT Bins. For each Range-FFT bin of interest, RadSee first adjusts its DC component to the dynamic range of its real and imaginary parts, and then applies a low-pass filter to remove the high-frequency component. As per [42], RadSee sets the low-pass filter's bandwidth to 5 Hz.

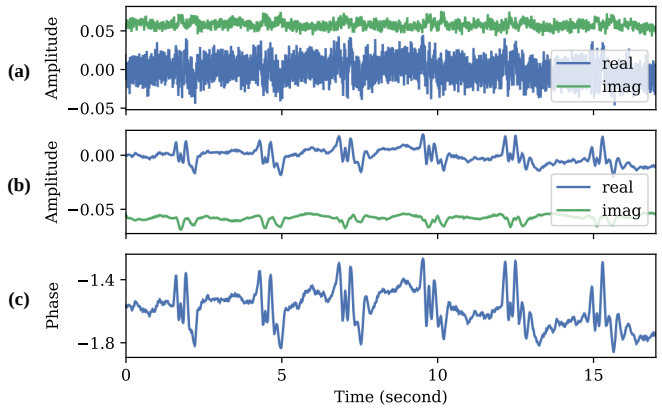


Fig. 11: (a) The original signal of one Range-FFT bin (one sample per chirp cycle); (b) the Range-FFT bin after DC adjustment and low-pass filter; (c) phase of the signal in (b).

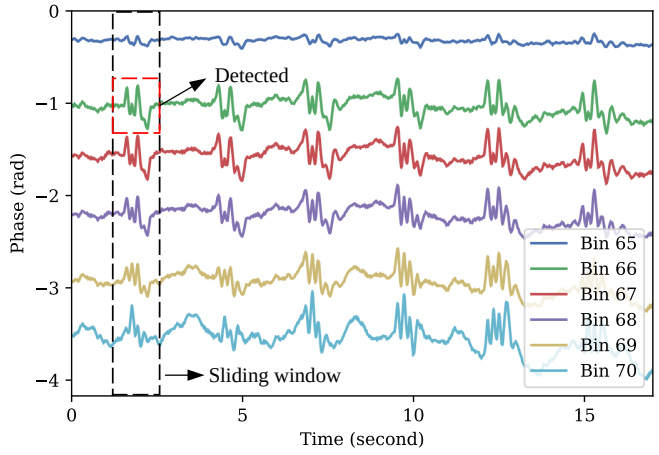


Fig. 12: Phase sequence of six Range-FFT bins.

Fig. 11 compares the data sequences of one Range-FFT bin *before* and *after* the DC adjustment and low-pass filter. It can be observed that the process can manifest the phase pattern of handwriting effectively.

FFT Bin Selection. Experiments show that handwriting will cause multiple bins to fluctuate. This can be attributed to the high range resolution and the multi-path effect within antenna's aperture. Instead of using a single Range-FFT bin, RadSee uses *multiple consecutive* Range-FFT bins to extract their phase patterns. The questions need to be answered: (i) how many Range-FFT bins should be selected, and (ii) which Range-FFT bins should be used. For the first question, RadSee empirically selects *five* consecutive Range-FFT bins and uses their phase information for letter classification. For the second question, RadSee selects the Range-FFT bins of the *smallest* index but with its phase variance larger than a predefined threshold. RadSee's bin selection algorithm is provided in Alg. 1. Its core idea is to identify five consecutive FFT-Range bins based on their phase variances, so that the handwriting movement pattern can be captured along the line-of-sight (shortest) through-wall path. These five bins are then fed into our DNN for letter recognition. Fig. 12 shows a sample of our observed Range-FFT bins in handwriting detection. In this case, RadSee selects bins 66 to 70 as the input of its DNN model for letter classification.

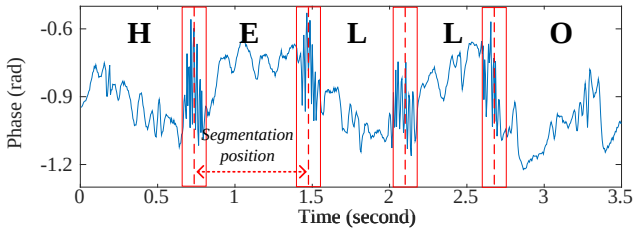


Fig. 13: Illustrating the rapid phase change of a target Range-FFT bin during the transition of writing letters.

Algorithm 1 RadSee's bin selection algorithm.

Input: Range-FFT phase matrix $[S(i,t) \in \mathbb{R}]_{N\times T}$, where i is bin index ($0 \le i < N$), t is time index ($0 \le t < T$), window size W, predefined lower bound of variance θ_{lw}, predefined upper bound of variance θ_{up}. ▷ *In our experiments,* $W = 500$, $N = 256$, $T = 5000$, $\theta_{lw} = 0.03$, $\theta_{up} = 0.18$.

Output: The smallest bin index i where the phase variance exceeds θ_{lw} but is lower than θ_{up}.

for $t = 0$ to $T - W$ **do**
  for $i = 0$ to N **do**
    Calculate window-slided variance as follows:
    $v(i,t) = \frac{1}{W}\Sigma_{j=t}^{t+W-1}|S(i,j) - \mu|^2$,
    where $\mu = \frac{1}{W}\Sigma_{j=t}^{t+W-1}S(i,j)$.
    if $v(i,t) > \theta_{lw}$ & $v(i,t) < \theta_{up}$ **then**
      return i
    end if
  end for
end for
return -1 ▷ *Indicate no writing activity is detected.*

Data Segmentation. RadSee performs data segmentation on the phase stream of the selected Range-FFT bins to extract the meaningful features that correspond to individual letters. RadSee employs different methods for phase data segmentation at the training and test phases. We elaborate them as follows. *(i) During Training Phase:* Since we have full control of the training data collection, we ask every participant to stop and be still for one second after writing each letter. By doing so, RadSee can easily segment phase sequence and extract meaningful phase data for individual letters. *(ii) During Test Phase:* In this phase, RadSee has no control over the writing style of a victim. Likely, the victim writes in a continuous manner without a stop in the middle. Interestingly, we always observed a rapid phase change during the transition from writing one letter to another. Fig. 13 shows an example of our observations. This is caused by the pen-holding hand's quick movement during the transition period. RadSee leverages this signature to segment the phase data

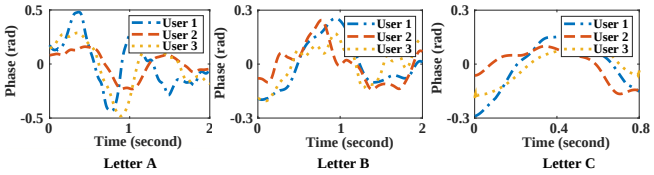


Fig. 14: The observed phase sequences when three users are writing letters 'A', 'B', and 'C'.

streams. Since the time duration of writing different letters may be different, the data sequences corresponding to different letters are of heterogeneous length.

Extracted Phase Features. Based on the above process, RadSee will obtain the phase data segments corresponding to individual letters being written. Fig. 14 shows some samples of its obtained phase segments from different users. From the figure we have the following observations. First, for the same user, the phase patterns of different letters are different. This is an encouraging observation as the uniqueness of phase patterns is the foundation of letter classification. Second, for the same letter (e.g., letter 'A' in Fig. 14), the phase patterns from different users look different. So far, it is not clear if those phase patterns will be classified to the same letter through an advanced transformation. To better understand this question, we conduct feature validation using kNN.

B. kNN-based Feature Validation

We use the kNN model [8] to validate the effectiveness of the extracted features. kNN is a simple data classification method that estimates the belonging of a new data sample based on a set of labeled data samples. When a new data sample comes, the distance between this new sample and all labeled samples is calculated. Then, the k closest neighbors are selected. The selected k closest neighbors cast weighted votes (using their distance) to make the final classification decision for the new data sample. One issue with kNN in this case is that the length of data samples (phase sequences) is not fixed, i.e., different phase sequences have different lengths. To address this issue, we employ Dynamic Time Warping (DTW), which has been widely used in speech recognition [10] and data mining [24]. DTW can find an optimal alignment between the two sequences by warping the time axis non-linearly.

Data Set. We collected the phase data samples for 62 letters (a-z, A-Z, and 0-9) from 12 users. Each user was asked to write in print writing style on a desk that is one meter away from the wall. Our radar was placed just behind the wall to collect the phase data. Each letter has 10 samples from a user and a total of 120 samples from those 12 users. In total, 7,440 samples were collected for all 62 letters, all of which were labeled during the data collection. The data samples are divided into two groups: those from the first 6 users are used for training, while those from the second 6 users are used for test.

Validation Results. We perform kNN on the collected data set. As an example, Fig. 15 shows the search results of kNN when the new data sample is the phase sequence of letter 'A'.

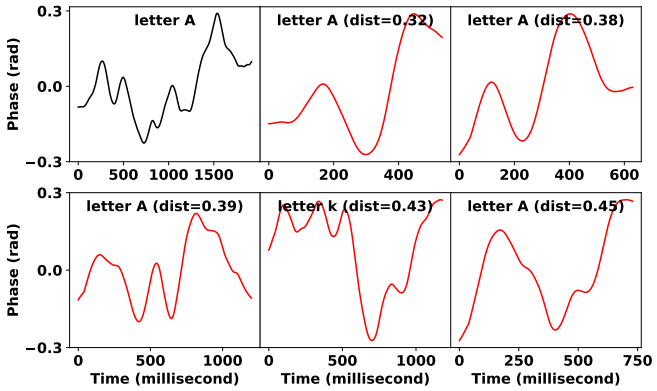


Fig. 15: Results of using kNN to search 5 closest neighbors for a new data sample. The top-left figure shows the phase sequence of the new data sample. The remaining 5 figures show the found 5 closest data samples (and their corresponding letters) in our training data set.

It can be seen that, of the five closest data samples in the training data set, four are correct (labeled with 'A') and one is incorrect (labeled with 'k'). The five closest data samples cast votes to make the final decision. The weighted vote for 'A' is 10.54, while the weighted vote for 'k' is 2.32. Based on the voting result, this new data sample is classified to letter 'A', which is correct.

Fig. 16 shows kNN's classification accuracy when the test data samples are from 6 different users. We note that the test data samples and the training data samples are from different users. As we can observe, the classification accuracy is from 53% (user 4) to 77% (user 3). This could be attributed to two factors: i) most of training data are from Asian participants; and ii) User 4 is an American participant while other five users are Asian participants.

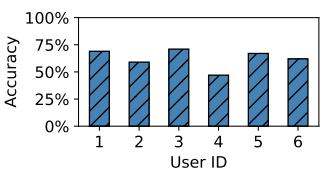


Fig. 16: kNN's classification accuracy when test and training data are from different users.

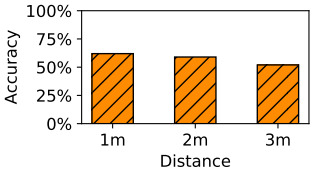


Fig. 17: kNN's classification accuracy when radar is at different distances.

We then evaluate kNN's classification accuracy using the data samples from User 6 when the radar was placed at different distances (1 m, 2 m, and 3 m). The training data samples were collected from six different users when the radar was placed at 1 m distance. Fig. 17 presents the classification results. It shows that the classification accuracy is 68% when the test was conducted at the same distance. However, when RadSee has a different distance from the victim, its detection accuracy decreases to 58%.

Limitations of kNN. The kNN-based classification results indeed manifest the effectiveness of phase features in

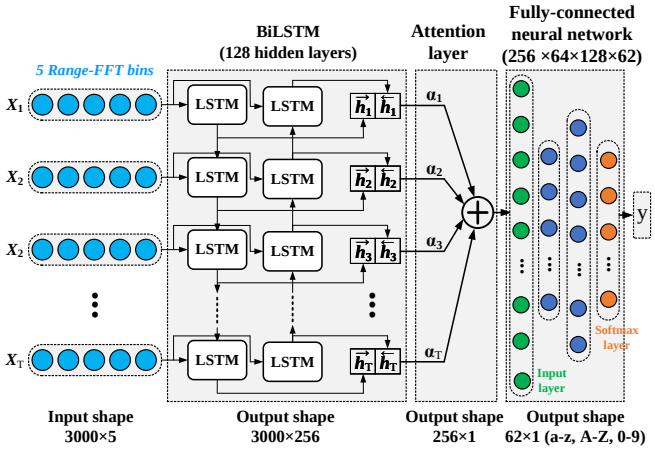


Fig. 18: The structure of an attention-based BiLSTM model for letter recognition. The input is the phases of selected 5 Range-FFT bins over $T = 3000$ ms, and the output is the classified one from the 62 characters.

handwriting letter classification. But this approach has two limitations. First, it has a very high computational complexity and thus limits the size of the labeled (training) data set. Second, it uses the phase sequence from only one Range-FFT bin for classification. Using those five Range-FFT bins together may improve the classification accuracy. In what follows, we design a DNN-based approach for handwriting recognition, with the aim of overcoming the above limitations and improving the classification accuracy.

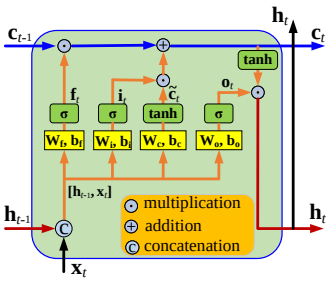
V. RadSee: DNN-based Recognition

In this section, we focus on designing a DNN model for through-wall handwriting recognition using the phase features extracted in the previous section. Compared to kNN, DNN is much more efficient in computation and is more appealing for practical use.

A. DNN Model

In essence, this letter recognition problem is a classification problem with its input being multi-dimensional phase sequences and its output being the probability of each letter in the candidate set (a-z, A-Z, and 0-9). We found that this task is similar to many classification tasks in natural language processing (NLP), such as information status classification [18] and stress detection [54]. Following the state-of-the-art classification techniques in NLP, we employ an attention-based Bidirectional LSTM (BiLSTM) model for RadSee's letter classification.

Fig. 18 shows the high-level structure of our attention-based BiLSTM model. The BiLSTM component is used to extract the temporal features in the time-series phase sequence. The attention layer is used to capture the key movement information of handwriting. This is critical as the key information of handwriting movement likely lies in some turning points. This attention layer will allow the model to focus on specific parts



$$\mathbf{f}_t = \sigma(\mathbf{W}_f[\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_f)$$
$$\mathbf{i}_t = \sigma(\mathbf{W}_i[\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_i)$$
$$\mathbf{o}_t = \sigma(\mathbf{W}_o[\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_o)$$
$$\tilde{\mathbf{c}}_t = \tanh(\mathbf{W}_c[\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_c)$$
$$\mathbf{c}_t = \mathbf{f}_j \odot \mathbf{c}_{t-1} + \mathbf{i}_t \odot \tilde{\mathbf{c}}_t$$
$$\mathbf{h}_t = \mathbf{o}_t \odot \tanh(\mathbf{c}_t),$$

Fig. 19: The structure and operation of an LSTM cell ($\mathbf{h}_t \in \mathbb{R}^{128\times 1}$, $\mathbf{c}_t \in \mathbb{R}^{128\times 1}$, and $\mathbf{W}_f, \mathbf{W}_i, \mathbf{W}_c, \mathbf{W}_o \in \mathbb{R}^{128\times 133}$).

(e.g., those turning points) of the phase sequences, thereby improving the accuracy and efficiency of classification.

B. BiLSTM

BiLSTM is a variant of the LSTM network [17] and has demonstrated its effectiveness for a wide range of NLP tasks such as machine translation [41], part-of-speech tagging [30], and sentiment analysis [49], [66]. In a BiLSTM, the input sequence is processed in both forward and backward directions using two separate LSTM layers. This allows the model to capture both past and future context for each input element. This is crucial for handwriting recognition, because the turning points of handwriting movement carry the key information for letter classification but the turning points may appear at the beginning, in the middle, and at the end of a phase sequence. The use of BiLSTM allows the model to capture those turning points at any pace of the input phase sequence.

Input Data. We set the input data shape to be 3000×5, where 3,000 is the number of chirps and 5 is the number of selected Range-FFT bins. Recall that each chirp is 1 ms. This means that the maximum time of writing a letter is 3 seconds. In most cases, one can finish the writing of a letter less than 3 seconds. If the phase sequence is less than 3,000 points, we simply pad zero behind the phase sequence as the input of BiLSTM. If the phase sequence is greater than 3,000, we trim the head and tail of the phase sequence, retaining only 3,000 points in the middle as input for the BiLSTM.

LSTM Cell. LSTM has been used in a wide range of learning tasks. It is the key component of the BiLSTM model as shown in Fig. 18. It allows the model to selectively retain or forget information at each time step. The cell structure includes three gates: an input gate, a forget gate, and an output gate. The input gate determines which information should be stored in the cell, the forget gate determines which information should be discarded, and the output gate determines which information should be used for the current output. Fig. 19 shows the structure and parameters of each LSTM cell.

BiLSTM Structure. As shown in Fig. 18, BiLSTM has two LSTM cells: one is for forward information flow, and the other is for backward information flow. In each iteration t, it combines the hidden states of forward and backward LSTMs through concatenation: $\mathbf{h}_t = [\vec{\mathbf{h}}_t, \overleftarrow{\mathbf{h}}_t]$, where $\vec{\mathbf{h}}_t$ is the hidden state from the forward LSTM, $\overleftarrow{\mathbf{h}}_t$ is the hidden

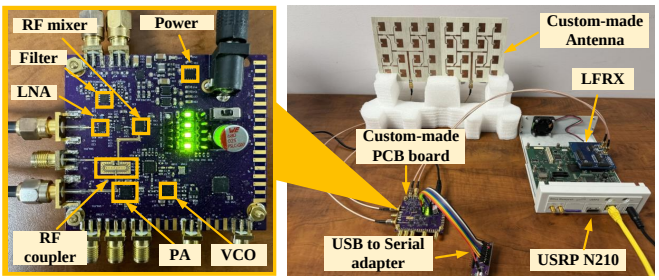


Fig. 20: Radar PCB (left) and a picture of RadSee (right).

state from the backward LSTM, and $\mathbf{h}_t$ is the hidden state of the BiLSTM. Since each LSTM has 128 hidden layers, we have $\mathbf{h}_t \in \mathbb{R}^{256\times 1}$, with $t = 1, 2, \ldots, 3000$. Then, the combined hidden states are fed to the attention layer for further processing.

C. Attention Layer

The attention mechanism is probably one of the most important inventions for deep learning and it has been used for many applications such as GPT [6], [39], [49], [60]. With the attention layer, the model learns to focus on some key parts of the data sequence. During the handwriting of a letter, some turning points may carry critical information for letter classification. The attention layer attempts to learn the importance of each part of the phase sequence and then assigns them with proper weights. To calculate the corresponding weights, it first feeds $\mathbf{h}_t$ to a one-layer Multilayer Perceptron (MLP) to learn a hidden representation u_t, and then normalizes the weights to generate α_t. Mathematically, it can be written as follows:

$$u_t = \tanh(\mathbf{W}_h^\top \mathbf{h}_t + b_h), \tag{5a}$$

$$\alpha_t = \frac{\exp(u_t)}{\sum_{k=1}^{T} \exp(u_k)}, \tag{5b}$$

$$\mathbf{s} = \sum_{t=1}^{T} \alpha_t \mathbf{h}_t, \tag{5c}$$

where $\mathbf{W}_h \in \mathbb{R}^{256\times 1}$ is the training weights, $b_h \in \mathbb{R}$ is a training bias, and $\mathbf{s} \in \mathbb{R}^{256\times 1}$ is the weighted vector for the fully-connected neural network in Fig. 18. The fully-connected network is of $256 \times 64 \times 128 \times 62$ size. The last layer is a SoftMax layer to calculate the possibility of each letter candidate (a-z, A-Z, and 0-9). The letter of the highest possibility is selected as the output y.

VI. Implementation

A. Hardware

Fig. 20 shows the hardware components of RadSee. We fabricated a radar PCB board as shown in this figure. The electronic components of this board include VCO, LNA, PA, Tx/Rx 16 dB RF coupler, RF quadrature mixer, and baseband filter. This PCB was made by OSH Park using FR408 substrate. We designed, simulated, and optimized 4×4 patch-array antennas using HFSS for radio signal transmission and reception. These antennas offer 18 dBi antenna gain for both transmission and reception. In total, it offers 36 dBi gain for the link path, making it possible to compensate the signal penetration loss of a wall. The total cost of RadSee is approximately $500, including $50 for PCB fabrication, $50 for antennas, and $400 for chips. We use USRP N210 with LFRX daughterboard to convert the analog signal to digital I/Q samples, which were then sent to a computer for data process. Transmission power is set to 20 dBm. The FMCW radar sweeps from 5.4 GHz to 6.5 GHz. The time duration of one chirp period is 1 ms, including 600 μs for frequency sweeping and 400 μs for idle.

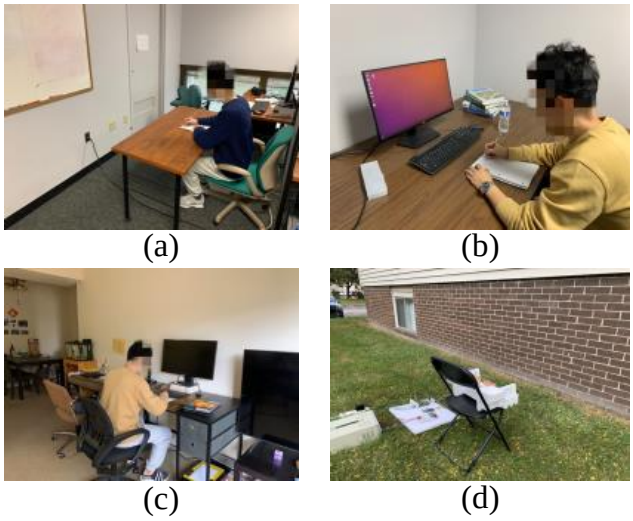


Fig. 21: Evaluation setting: (a) Laboratory scenario. (b) Office scenario. (c) Apartment scenario. (d) RadSee attacks from outside of the apartment.

B. Algorithms

Digital Signal Processing. We implemented the data processing algorithms on a laptop in C++ using GNU Radio Out-of-Tree (OOT) module. The laptop receives a continuous data stream from the radar. It needs to synchronize the chirp signal and extract the useful data samples of each chirp. Fortunately, due to the presence of 400 μs idle period of each chirp, it is easy to identify the useful data samples from the data stream. Specifically, we use the high peaks as shown in Fig. 10 to extract the useful data samples. One fundamental issue with the current hardware design is the lack of clock synchronization between ADC and FMCW chirps. To address this issue, we use a high sampling rate 5 MSps and perform fine-grained synchronization to identify the first data sample corresponding to the starting moment of each chirp.

Data Collection for DNN Training.[1] We collected training data in a laboratory. The radar was placed behind an interior

[1]The experiments do not require IRB approval based on the determination results from the authors' institution. The experiments were conducted under FCC experimental spectrum license with Call Sign # WM2XWQ and File # 0954-EX-CN-2022.

TABLE II: Participants for training and test data collection.

	Participants for training				Participants for test			
Handedness	Right-handed		Left-handed		Right-handed		Left-handed	
Writing style	Print	Cursive	Print	Cursive	Print	Cursive	Print	Cursive
# of participants	11	5	1	1	7	3	1	1

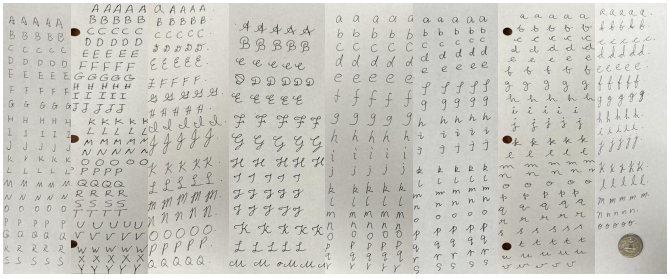


Fig. 22: Writing samples from participants for training.

drywall at a distance of 0.5 m. A writing desk was placed in front of the wall at a distance of 1 m, as shown in Fig. 21(a). Eighteen participants (4 American, 3 Indian, 4 Middle East, 7 Chinese) were asked to write 62 characters (a-z, A-Z, and 0-9) on the desk. Each participant wrote every character 60 times. In total, we collected $18 \times 62 \times 60 = 66,960$ data samples. Of the eighteen participants, twelve were asked to write in the *print* style, while six were asked to write in the *cursive* style. Regarding handedness, two of them were left-handed writers while the rest were right-handed writers. The handedness and writing styles of the participants are summarized in Table II. Some writing samples from the participants are provided in Fig. 22.

DNN Training. The DNN model was implemented using TensorFlow's Keras library. We used cross entropy as loss function. During the training process, we set the batch size to 2,000 and trained the model for 500 epochs. We used Adam optimizer with a learning rate of $7e^{-4}$ to train the model.

VII. Experimental Evaluation

A. Letter Recognition Accuracy

Write on A4 Papers. Recall that our training data was collected in a laboratory from eighteen participants. To evaluate the recognition accuracy of RadSee, we completely separate the training and test datasets. We invited twelve new participants (4 American, 4 Chinese, 2 Indian, 2 Middle East) to write letters in the same setting (i.e., sitting 1 m away from the wall and facing to the radar). None of these twelve people participated in the training data collection. Each of them wrote 300 random letters on A4 papers. During the test, eight participants were asked to write in the *print* style, and four were asked to write in the *cursive* style. Both print and cursive writing letters are within the size of 5 mm to 10 mm. Regarding handedness, ten participants were right-handed writers, while two were left-handed writers. The handedness and writing style are summarized in Table II.

Fig. 23 shows the confusion matrix of RadSee's letter recognition results. It is evident that RadSee can recognize most of the letters. RadSee is prone to making mistakes for some letters. For instance, it can easily confuse 'O' with 'o',

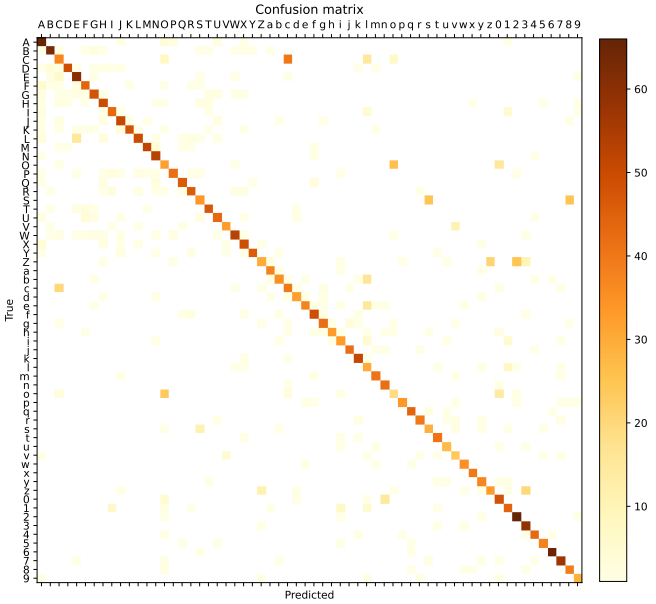


Fig. 23: Confusion matrix of RadSee's letter recognition results.

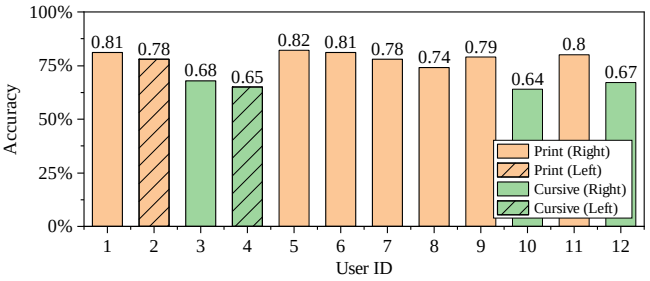


Fig. 24: RadSee's letter recognition accuracy when participants wrote on A4 papers. Users 1-4 are Americans, users 5-8 are Chinese, users 9-10 are Indians, and users 11-12 are from Middle East.

'C' with 'O', and 'I' with '1'. Other errors can arise from cursive writing, such as confusing 'S' with '8' and 'Z' with '3'. This is understandable, as their handwriting patterns are similar to each other. Overall, RadSee achieves 75% letter recognition accuracy.

Print vs. Cursive. Fig. 24 presents RadSee's letter recognition accuracy for the 12 individual participants. As observed, RadSee has a lower recognition accuracy for the participants who wrote in cursive style compared to those who wrote in print style. This observation can be attributed to two factors. First, cursive writing is more individualized and diverse, making it challenging for the model to extract consistent features across different participants, despite having cursive-style data in the training dataset. Second, our segmentation method relies on detecting signal transitions between letters, which becomes more difficult when people write in cursive style.

Writing Handedness. Besides writing style, handedness is another factor that may affect RadSee's letter recognition

(a) (b)

Fig. 25: Writing on different media. (a) Writing on papers, iPad, and Post-it notes. (b) The recognition accuracy of RadSee when writing on different media.

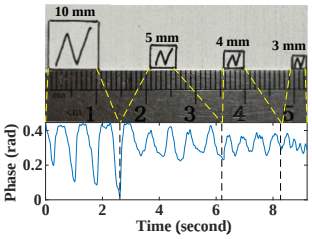


Fig. 26: RadSee's phase signal for different letter size.

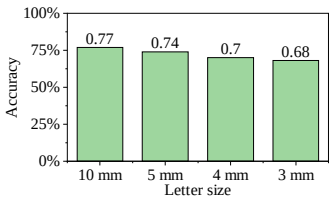


Fig. 27: RadSee's accuracy for letters of different sizes.

accuracy. However, experimental results show that handedness affects RadSee very slightly. As shown in Fig. 24, RadSee has a very similar performance for both left-handed and right-handed users. This can be attributed to the fact that most left-handed individuals have the same writing movement pattern as right-handed individuals, i.e., write from left to right and from top to bottom.

Write on iPad and Post-it Notes. Tablets, such as Apple iPad, have become increasingly popular for writing activities, with many individuals opting to use them for important documents instead of traditional pen and paper. To evaluate the performance of writing on an iPad, we repeated our measurements by asking twelve participants to write 300 random letters using an Apple Pencil. The experimental results are shown in Fig. 25(b). RadSee achieves 74% letter recognition accuracy. In the same setting, RadSee achieves 75% letter recognition accuracy when participants write on A4 papers. This indicates that RadSee has almost the same performance for A4 paper and iPad writing recognition. Another commonly used medium for writing is Post-it notes. Given their smaller size, we asked participants to write 20 random letters on Post-it notes. RadSee's letter recognition accuracy for Post-it notes is 71%, as presented in Fig. 25(b). As shown in Fig. 25(a), these three writing media have different horizontal writing ranges. Since RadSee has similar performance for them, it suggests that RadSee effectively accommodates the horizontal range for writing on A4 papers, iPad, or Post-it notes.

B. Impact of Letter Size

We conducted experiments to better understand RadSee's ability of detecting small-size letters. Fig. 26 presents RadSee's signal changes when a participant wrote letter 'N' of different sizes. Evidently, RadSee is capable of detecting as small as 3 mm handwriting movement. We further asked one participant to write on A4 papers with grid boxes of different sizes: 3 mm $\times$ 3 mm, 4 mm $\times$ 4 mm, 5 mm $\times$ 5 mm, and 10 mm $\times$ 10 mm. The participant was instructed to write letters within the boundaries of the grid boxes. However, for the 3 mm $\times$ 3 mm grids, since the boxes were too small, a considerable portion of the written letters exceeded the boundaries. Fig. 27 presents RadSee's letter recognition accuracy in these four cases. It is evident that RadSee's accuracy decreases with the letter size. But notably, RadSee achieves 68% recognition accuracy even in the case where the letter size is confined within 3 mm.

C. Impacts of Distance and Angle

When an attacker attempts to detect the handwriting behind a wall, it may not know the distance from itself to the victim and the angular direction of the victim. The attacker may use RadSee to do an exhaustive search to find the best pointing direction for the radar's antennas, but the radar-antenna-pointing direction may not be accurate. To evaluate RadSee's robustness, we examine its accuracy in different settings: (i) the writers are 1 m, 2 m, and 3 m behind the wall; and (ii) RadSee's antenna is pointing to different angles (0°, 10°, 20°, and 30°). The combination constitutes 12 different cases. In each case, we instructed eight participants to write 300 letters using their normal handwriting habits.

Fig. 28 presents our measured accuracy and deviation. It can be seen that RadSee is robust to the distance change. This can be explained by its design. In nature, FMCW radar is capable of precisely capturing the movement features at different distances. When the distance between the writer and the wall changes from 1 m to 3 m, RadSee will identify another 5 Range-FFT bins for phase feature extraction. Since the handwriting movement patterns are not related to the wall distance, the extracted features will remain unchanged. Therefore, RadSee is robust to distance changes.

Fig. 28 also presents our measurement results when RadSee's antennas was pointing to different angles. Evidently, RadSee's accuracy decreases when its directional error increases from 0° to 30°. Specifically, when RadSee was pointing to 0°, it achieved 77% recognition accuracy. When RadSee was pointing to 30°, it achieved 55% recognition accuracy. In all cases, the standard deviation is almost the same, i.e., 4%. This degradation can be attributed to the directivity of the patch-array antennas, as shown in Fig. 6. When the writer deviates from its central direction, the patch antenna's effective radiation power decreases, making noise and other imperfections more significant and thus leading to a decreased accuracy.

D. Impact of Interference from Other Moving Objects

Experimental results in Fig. 7 have confirmed that RadSee is immune to radio interference from in-band (5 GHz) Wi-Fi devices. All experiments in this work were conducted in office and laboratory environments, which are rich with interference from multiple Wi-Fi sources. Therefore, the experimental results presented have already taken into account the

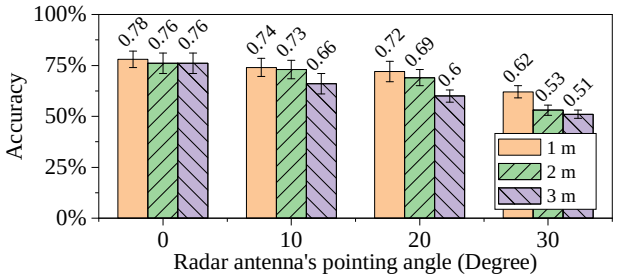
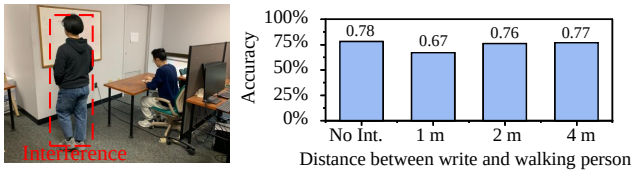


Fig. 28: Letter recognition accuracy of RadSee when writers are at different distances and different angles from the wall.



(a) (b)

Fig. 29: Interference test. (a) Interferer is 2 meters from writer. (b) RadSee's resilience to interference from a walking person.

Fig. 30: Illustration of six different types of wall materials.

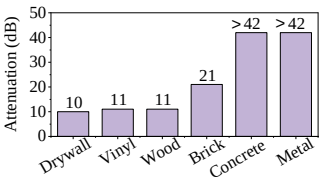


Fig. 31: RF signal's power attenuation for penetrating a wall of different materials.

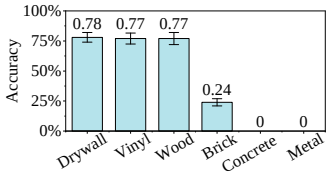


Fig. 32: RadSee's recognition accuracy when placed behind six wall materials.

radio interference from multiple Wi-Fi sources. Additionally, RadSee is not affected by static objects (e.g., desks and chairs) around a writer as they appear to be a constant in the received signal, which can be easily mitigated. Therefore, we focus on studying RadSee's performance in the presence of moving objects (e.g., walking persons) in the proximity of the writer. We emulated this scenario by asking another person to walk around the writer as shown in Fig. 29(a). We measure the recognition accuracy of RadSee in three cases, i.e., the distance between a writer and a walking person is 1 m, 2 m, and 4 m. We asked eight participants to write 300 random letters in each case and measured RadSee's letter recognition accuracy.

Fig. 29(b) depicts our measured results. We can see that the performance degradation depends on the distance between the writer and the interferer. The closer the interferer is, the larger performance degradation RadSee has. For the case where interferer is 1 m away, RadSee demonstrates 67% letter recognition accuracy, with 11% accuracy degradation compared to the case without interference. When the interferer is 2 m away, RadSee rapidly increases its accuracy to 76%, which is close to its accuracy in the case without interference. We note that the participants in all experiments maintained normal physiological activities, such as breathing and respiration. The experimental results reported above have already taken into account those normal physiological activities of the writers.

E. Impact of Different Wall Materials

RF signals have varying penetration abilities depending on the type of wall. We conducted experiments to evaluate the performance of RadSee in detecting letters through different wall materials. Specifically, we considered six wall materials as shown in Fig. 30: drywall (12 cm), vinyl wall (20 cm), wood wall (19 cm), brick wall (22 cm), concrete wall (23 cm), and metal door (4 cm). We first measured their penetration loss, which refers to the power attenuation of radio signals as they pass through a wall. Fig. 31 presents our measurement results. It is evident that drywall, vinyl and wood walls have similar penetration loss for radio signal, which is about 10 dB. Brick wall is more lossy for radio signal compared to wood wall. Its penetration loss is about 21 dB. However, concrete walls and metal doors completely block radio signals. Their attenuation loss is greater than 42 dB.

We then conducted experiments to measure RadSee's letter recognition accuracy. Eight participants took part in the experiments. They were seated 1 meter away from the wall, while RadSee was positioned 0.5 meters away on the other side of the wall as shown in Fig. 21. Each of the eight participants wrote 300 random letters using his/her own writing style. Fig. 32 presents the experimental results. It shows that RadSee achieves similar performance when participants wrote behind drywall, vinyl, and wood walls. This similarity is due to the comparable electromagnetic properties of these materials. In contrast, a brick wall significantly reduces recognition accuracy, with RadSee achieving only 24% letter recognition accuracy in this scenario. Furthermore, concrete walls and metal doors completely obstruct letter detection.

F. Word Recognition Accuracy in Content

In addition to detecting individual letters, we evaluate RadSee's performance of recovering entire sentences. This is important because an attacker's interest may lies in the content that a victim is writing, rather than individual letters. We asked twelve participants to reproduce an CNN News article, which is about 300 words. Some writing samples from the

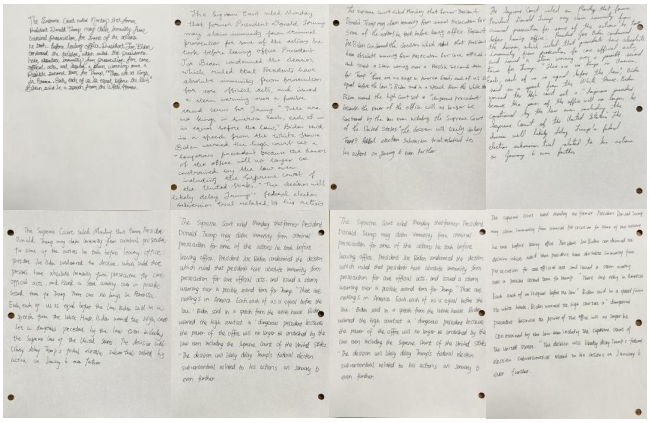


Fig. 33: Writing samples from participants as they transcribed CNN news articles in both print and cursive styles.

TABLE III: A case study of RadSee detecting the sentences written by a person behind a lab drywall.

Ground truth	Letters recognized by RadSee	Segmented by `Wordsegment` [20]	Corrected by `TextBlob` [34]
'football is popular in the united states'	'ecctbollispo pulaintheuni tedstate'	'ecc', 't', 'boll', 'is', 'popula', 'in', 'the', 'united', 'state'	'etc', 't', 'ball', 'is', 'popular', 'in', 'the', 'united', 'state'
'Bill is a hardworking student'	'Billiislhar dworkimg studena'	'bill', 'i', 'isl', 'hard', 'work', 'img', 'studena'	'bill', 'is', 'hard', 'work', 'ing', 'student'
'My favourite fruit is apple'	'mgfavouri teffruitl qapple'	'mg', 'favourite', 'f', 'fruit', 'lq', 'apple'	'my', 'favourite', 'fruit', 'is', 'apple'

participants are provided in Fig. 33. The experimental setting is the same as described above.

RadSee employs two open-source software tools to translate its detected letters into word sentences: `Wordsegment` [20] and `TextBlob` [34]. It first sends the detected letters to `Wordsegment` for word segmentation. Then, it sends the segmented text to `TextBlob` for automatic spelling correction. Table III presents samples of the sentence recognition results. Leveraging these two open-source tools, RadSee demonstrates impressive performance in word and sentence recognition. It nearly recognized the first sentence in the table and accurately recovered both the second and third sentences.

We then use *word recognition accuracy* as the metric to evaluate the performance of RadSee. According to [35], word recognition accuracy is defined as $WRA = \frac{N-S-D-I}{N}$, where N is the number of words in the ground-truth text, S is the number of word substitutions, D is the number of word deletions, and I is the number of word insertions. Fig. 34 shows RadSee's *WRA* with and without using `TextBlob` for automatic spelling correction. It can be seen that without automatic spelling correction, RadSee's *WRA* ranges from 40% to 56% across the twelve participants. In contrast, when automatic spelling correction is applied, RadSee's *WRA* significantly improves, ranging from 79% to 93%. On average, RadSee's *WRA* hovers around 87% with automatic spelling correction. This level of word recognition accuracy is sufficient for an attacker to comprehend the content written by a victim.

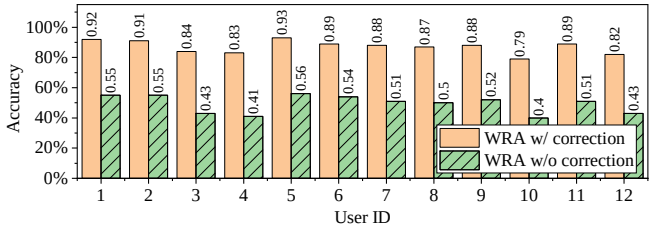


Fig. 34: Word recognition accuracy of RadSee with and without correction for different users.

VIII. Countermeasures and Other Applications

A. Countermeasures

Handwriting Safety Tips. RadSee demonstrated a serious threat to handwriting privacy. Based on the study, we have the following tips for those who have concerns about their handwriting information leakage. **Tip 1:** Do not write important documents in a room with drywall or vinyl wall. Instead, write them in a room with thick concrete or any metal walls. These walls can largely reduce the radio signal and thus reduce the probability of information leakage. **Tip 2:** Do not face yourself to a wall behind which a radar may be placed. Instead, face against that wall. Your body/torso will significantly reduce the radio signal strength and thus reduce the probability of your content being detected by an attacker. **Tip 3:** If possible, write important documents on a desk far from all walls rather than a desk against a wall. This will increase the distance between yourself and a radar, thereby reducing its recognition accuracy.

Protection Strategies. One natural approach to protecting handwriting content is to install multi-layer RF shielding materials inside the walls of your room [26]. Common materials used for RF shielding include metals such as aluminum, copper, and steel, as well as conductive coatings or paints. Another approach is to take advantage of recent advances in reconfigurable intelligent surface (RIS), which has also been studied under other names such as electromagnetic metasurface or radio relay. RIS can be used to create virtual multipath from radar's Tx to its Rx. By manipulating its phase shifting and beam steering, RIS is capable of generating fake phase patterns for the radar, preventing it from recovering the handwriting content. Unfortunately, neither of the above approaches is easy or economical to deploy.

B. Other Applications

While RadSee was designed to better understand the radio attacks related to handwriting privacy, it can also be used for many other applications. For instance, RadSee can be installed on a laptop as an input method. When an end user physically writes something on paper in front of his/her laptop, the content is automatically recognized by RadSee and digitally recorded on his/her laptop. In this case, RadSee does not need to use a 4×4 patch-array antennas since there is no need to penetrate through walls. Rather, a small patch antenna should be sufficient. RadSee can also be used as a human-computer interface for smart TVs. End users can write using their bare

hands, and a TV equipped with RadSee can recognize the letters being written.

IX. Related Work

We surveyed the literature in two categories: *through-wall detection* and *fine-grained human activity recognition*. Table I in Section I outlined RadSee's uniqueness compared to prior work.

A. See Through Wall using Radio

See Through Wall using FMCW Radar. Some pioneering works have studied 6 GHz FMCW radar to detect and track human activities behind walls using model-based or learning-based methods [1], [29], [63]–[65]. For instance, [63]–[65] focuses on using FMCW radar to generate the heatmap image of human body skeleton through walls. [29] uses FMCW radar to detect the interactions between two people behind walls. However, all these works are based on the ranging detection of FMCW radars. Since the range resolution of an FMCW radar is fundamentally limited by its bandwidth, this method cannot achieve mm-level accuracy for through-wall motion detection. To address this issue, RadSee uses the phase information for through-wall mm-level hand movement detection.

RF-capture [1] is probably the most related work of RadSee. It also uses FMCW radar to recognize the "handwriting" behind a wall. However, the letters that RF-capture aims to recognize are of large size (e.g., 0.5 m$\times$0.5 m). It is actually a gesture recognition rather than normal-sized handwriting detection. Its method is based on range- and angle-based tracking, and thus cannot achieve mm-level accuracy. Therefore, RadSee is fundamentally different from RF-capture.

Through-Wall Detection using Wi-Fi. Wi-Fi signal is ubiquitous and it has a strong ability of passing through a wall. [2] utilizes Wi-Fi signals and multi-antenna techniques to track the movement of people behind a wall. [52] uses Wi-Fi signals to recover the audio sound from a speaker placed behind a soundproof wall. However, due to the no-coherent detection at a Wi-Fi receiver, it is impossible for a Wi-Fi receiver to detect movement at the millimeter level. Therefore, Wi-Fi signals are not suitable for through-wall handwriting detection.

Through-Wall Detection using RFID. Through-wall detection is also possible by using RFID systems. Tadar [59] and RF-HMS [51] demonstrated their capabilities of tracking human moving directions through walls using an array of RFID tags. However, the tracking error in these systems is around 10 cm, indicating their incapability of tracking mm-level hand movements. RFID tag can also be used to measure the vibration pattern of a loudspeaker [44]. But, due to its long wavelength (33 cm), it is not a good candidate for tracking mm-level movements.

B. Fine-Grained HAR

Handwriting Recognition. Camera-based handwriting recognition is a well-established field [7]. However, the camera cannot see through walls. Recently, RF signals have been studied for handwriting recognition. RF-IDraw [46] attaches an RFID tag to a people's finger and can reconstruct the trajectory of that finger. A multi-resolution positioning technique was designed, yielding a tracing accuracy at the centimeter level. mTrack [53] developed a mmWave (60 GHz) tracking system and achieved mm-level tracking accuracy. It also demonstrated its capability of recognizing handwriting letters. However, mmWave signals are vulnerable to blockage and cannot go through walls. Therefore, it is not suitable for our purpose.

MmWave FMCW Radar Detection. In recent years, mmWave (24 GHz, 60 GHz and 77 GHz) FMCW radars become available on the market for autonomous driving applications. These radars have been widely used for human activity recognition and vital sign detection [4], [15], [19], [31], [32], [43], [53], [57], [62]. Given their large bandwidth and small wavelength, they can easily achieve mm-level accuracy when detecting object movements. However, mmWave signals cannot pass through walls. Therefore, they cannot apply to through-wall handwriting detection.

Gesture and Vital Sign Detection. CSI in Wi-Fi networks has been used for a wide range of sensing applications such as gesture recognition [13], [27], [38], vital sign detection [48], and radio imaging [21], [28], [45]. However, Wi-Fi is a non-coherent system due to the physical separation of its transmitter and receiver. Therefore, its detection accuracy is fundamentally limited by timing, frequency, and phase misalignments. As a result, it is not competent for mm-level handwriting detection.

X. Conclusion

While mmWave FMCW radar has been extensively studied for autonomous driving and HAR, sub-10GHz FMCW radar has not received as much attention. This is of particular interest due to its *see-through-wall* capability, which may pose significant threats to the privacy of human activities. In this work, we presented RadSee, a 6 GHz FMCW radar system designed for detecting handwriting content behind walls. Through a combined hardware and software design, RadSee is capable of detecting mm-level handwriting movements and recognizing most letters based on their unique phase patterns. Additionally, it is resilient to the interference from other moving objects and coexisting radio sources. Extensive experimental results show that RadSee achieves 75% letter recognition accuracy when victims write 62 different letters and 87% word recognition accuracy when they write articles. In light of these realistic threats, we offered handwriting safety tips and defense strategies to help the public protect their handwriting information.

Acknowledgments

We sincerely thank the anonymous reviewers and our shepherd for their insightful comments. This project was supported in part by NSF Grants ECCS-2225337 and CNS-2100112.

References

[1] F. Adib, C.-Y. Hsu, H. Mao, D. Katabi, and F. Durand, "Capturing the human figure through a wall," *ACM Transactions on Graphics (TOG)*, vol. 34, no. 6, pp. 1–13, 2015.

[2] F. Adib and D. Katabi, "See through walls with wifi!" in *Proceedings of the ACM SIGCOMM 2013 conference on SIGCOMM*, 2013, pp. 75–86.

[3] AFPRELAXNEWS, "Handwriting still has a place in our connected world, now it's a trend on social media," https://tinyurl.com/bddxy57n, April 2024, [Online; accessed 01-April-2024].

[4] A. Ahmad, J. C. Roh, D. Wang, and A. Dubey, "Vital signs monitoring of multiple people using a fmcw millimeter-wave sensor," in *2018 IEEE Radar Conference (RadarConf18)*. IEEE, 2018, pp. 1450–1455.

[5] K. Ali, A. X. Liu, W. Wang, and M. Shahzad, "Keystroke recognition using wifi signals," in *Proceedings of the 21st annual international conference on mobile computing and networking*, 2015, pp. 90–102.

[6] C. Baziotis, N. Pelekis, and C. Doulkeridis, "Datastories at semeval-2017 task 4: Deep lstm with attention for message-level and topic-based sentiment analysis," in *Proceedings of the 11th international workshop on semantic evaluation (SemEval-2017)*, 2017, pp. 747–754.

[7] U. Choudhary, S. Bhosale, S. Bhise, and P. Chilveri, "A survey: Cursive handwriting recognition techniques," in *2017 2nd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT)*. IEEE, 2017, pp. 1712–1716.

[8] T. Cover and P. Hart, "Nearest neighbor pattern classification," *IEEE transactions on information theory*, vol. 13, no. 1, pp. 21–27, 1967.

[9] Dassault Systèmes Simulia], "CST Studio Suite," https://www.3ds.com/products-services/simulia/products/cst-studio-suite/, April 2024, [Online; accessed 17-April-2024].

[10] S. D. Dhingra, G. Nijhawan, and P. Pandit, "Isolated speech recognition using mfcc and dtw," *International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering*, vol. 2, no. 8, pp. 4085–4092, 2013.

[11] M. Ehatisham-Ul-Haq, A. Javed, M. A. Azam, H. M. Malik, A. Irtaza, I. H. Lee, and M. T. Mahmood, "Robust human activity recognition using multimodal feature-level fusion," *IEEE Access*, vol. 7, pp. 60 736–60 751, 2019.

[12] Y. Fu, S. Wang, L. Zhong, L. Chen, J. Ren, and Y. Zhang, "Svoice: Enabling voice communication in silence via acoustic sensing on commodity devices," in *Proceedings of the 20th ACM Conference on Embedded Networked Sensor Systems*, 2022, pp. 622–636.

[13] Z. Fu, J. Xu, Z. Zhu, A. X. Liu, and X. Sun, "Writing in the air with wifi signals for virtual reality devices," *IEEE Transactions on Mobile Computing*, vol. 18, no. 2, pp. 473–484, 2018.

[14] A. Graves, M. Liwicki, S. Fernández, R. Bertolami, H. Bunke, and J. Schmidhuber, "A novel connectionist system for unconstrained handwriting recognition," *IEEE transactions on pattern analysis and machine intelligence*, vol. 31, no. 5, pp. 855–868, 2008.

[15] U. Ha, S. Assana, and F. Adib, "Contactless seismocardiography via deep learning radars," in *Proceedings of the 26th annual international conference on mobile computing and networking*, 2020, pp. 1–14.

[16] W. He, K. Wu, Y. Zou, and Z. Ming, "Wig: Wifi-based gesture recognition system," in *2015 24th International Conference on Computer Communication and Networks (ICCCN)*. IEEE, 2015, pp. 1–7.

[17] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[18] Y. Hou, "Incremental fine-grained information status classification using attention-based lstms," in *Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers*, 2016, pp. 1880–1890.

[19] P. Hu, Y. Ma, P. S. Santhalingam, P. H. Pathak, and X. Cheng, "Milliear: Millimeter-wave acoustic eavesdropping with unconstrained vocabulary," in *IEEE INFOCOM 2022-IEEE Conference on Computer Communications*. IEEE, 2022, pp. 11–20.

[20] G. Jenks, "wordsegment," https://grantjenks.com/docs/wordsegment/, April 2024, [Online; accessed 17-April-2024].

[21] W. Jiang, H. Xue, C. Miao, S. Wang, S. Lin, C. Tian, S. Murali, H. Hu, Z. Sun, and L. Su, "Towards 3d human pose construction using wifi," in *Proceedings of the 26th Annual International Conference on Mobile Computing and Networking*, 2020, pp. 1–14.

[22] S.-R. Ke, H. L. U. Thuc, Y.-J. Lee, J.-N. Hwang, J.-H. Yoo, and K.-H. Choi, "A review on video-based human activity recognition," *Computers*, vol. 2, no. 2, pp. 88–131, 2013.

[23] B. Kellogg, V. Talla, and S. Gollakota, "Bringing gesture recognition to all devices," in *11th USENIX Symposium on Networked Systems Design and Implementation (NSDI 14)*, 2014, pp. 303–316.

[24] E. J. Keogh and M. J. Pazzani, "Scaling up dynamic time warping for datamining applications," in *Proceedings of the sixth ACM SIGKDD international conference on Knowledge discovery and data mining*, 2000, pp. 285–289.

[25] M. Kotaru, K. Joshi, D. Bharadia, and S. Katti, "Spotfi: Decimeter level localization using wifi," in *Proceedings of the 2015 ACM Conference on Special Interest Group on Data Communication*, 2015, pp. 269–282.

[26] B. J. Lee, R. D. Watkins, C.-M. Chang, and C. S. Levin, "Low eddy current rf shielding enclosure designs for 3t mr applications," *Magnetic resonance in medicine*, vol. 79, no. 3, pp. 1745–1752, 2018.

[27] C. Li, M. Liu, and Z. Cao, "Wihf: Enable user identified gesture recognition with wifi," in *IEEE INFOCOM 2020-IEEE Conference on Computer Communications*. IEEE, 2020, pp. 586–595.

[28] C. Li, Z. Liu, Y. Yao, Z. Cao, M. Zhang, and Y. Liu, "Wi-fi see it all: generative adversarial network-augmented versatile wi-fi imaging," in *Proceedings of the 18th Conference on Embedded Networked Sensor Systems*, 2020, pp. 436–448.

[29] T. Li, L. Fan, M. Zhao, Y. Liu, and D. Katabi, "Making the invisible visible: Action recognition through walls and occlusions," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2019, pp. 872–881.

[30] Y. Li, T. Liu, J. Jiang, and L. Zhang, "Hashtag recommendation with topical attention-based lstm," in *Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers*, 2016, pp. 3019–3029.

[31] Z. Li, F. Ma, A. S. Rathore, Z. Yang, B. Chen, L. Su, and W. Xu, "Wavespy: Remote and through-wall screen attack via mmwave sensing," in *2020 IEEE Symposium on Security and Privacy (SP)*. IEEE, 2020, pp. 217–232.

[32] J. Lien, N. Gillian, M. E. Karagozler, P. Amihood, C. Schwesig, E. Olson, H. Raja, and I. Poupyrev, "Soli: Ubiquitous gesture sensing with millimeter wave radar," *ACM Transactions on Graphics (TOG)*, vol. 35, no. 4, pp. 1–19, 2016.

[33] W. Lin, M.-T. Sun, R. Poovandran, and Z. Zhang, "Human activity recognition for video surveillance," in *2008 IEEE International Symposium on Circuits and Systems (ISCAS)*. IEEE, 2008, pp. 2737–2740.

[34] S. Loria, "TextBlob: Simplified Text Processing," https://textblob.readthedocs.io/en/dev/, April 2024, [Online; accessed 17-April-2024].

[35] I. McCowan, D. Moore, J. Dines, D. Gatica-Perez, M. Flynn, P. Wellner, and H. Bourlard, "On the use of information retrieval measures for speech recognition," tech. rep., IDIAP Research Institute, Martigny, Switzerland, Tech. Rep., 2005.

[36] J. McIntosh, A. Marzo, M. Fraser, and C. Phillips, "Echoflex: Hand gesture recognition using ultrasound imaging," in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 2017, pp. 1923–1934.

[37] K. Niu, F. Zhang, J. Xiong, X. Li, E. Yi, and D. Zhang, "Boosting fine-grained activity sensing by embracing wireless multipath effects," in *Proceedings of the 14th International Conference on emerging Networking EXperiments and Technologies*, 2018, pp. 139–151.

[38] Q. Pu, S. Gupta, S. Gollakota, and S. Patel, "Whole-home gesture recognition using wireless signals," in *Proceedings of the 19th annual international conference on Mobile computing & networking*, 2013, pp. 27–38.

[39] C. Raffel and D. P. Ellis, "Feed-forward networks with attention can solve some long-term memory problems," *arXiv preprint arXiv:1512.08756*, 2015.

[40] P. Roy, S. Ghosh, and U. Pal, "A cnn based framework for unistroke numeral recognition in air-writing," in *2018 16th international conference on frontiers in handwriting recognition (ICFHR)*. IEEE, 2018, pp. 404–409.

[41] H. Tao, S. Tong, H. Zhao, T. Xu, B. Jin, and Q. Liu, "A radical-aware attention-based model for chinese text classification," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 33, 2019, pp. 5125–5132.

[42] H.-L. H. Teulings and A. J. Thomassen, "Computer-aided analysis of handwriting movements," *Visible Language*, vol. 13, no. 3, p. 218, 1979.

[43] C. Wang, F. Lin, Z. Ba, F. Zhang, W. Xu, and K. Ren, "Wavesdropper: Through-wall word detection of human speech via commercial mmwave devices," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 6, no. 2, pp. 1–26, 2022.

[44] C. Wang, L. Xie, Y. Lin, W. Wang, Y. Chen, Y. Bu, K. Zhang, and S. Lu, "Thru-the-wall eavesdropping on loudspeakers via rfid by capturing sub-mm level vibration," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 5, no. 4, pp. 1–25, 2021.

[45] F. Wang, S. Zhou, S. Panev, J. Han, and D. Huang, "Person-in-wifi: Fine-grained person perception using wifi," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2019, pp. 5452–5461.

[46] J. Wang, D. Vasisht, and D. Katabi, "Rf-idraw: Virtual touch screen in the air using rf signals," *ACM SIGCOMM Computer Communication Review*, vol. 44, no. 4, pp. 235–246, 2014.

[47] S. Wang, J. Song, J. Lien, I. Poupyrev, and O. Hilliges, "Interacting with soli: Exploring fine-grained dynamic gesture recognition in the radio-frequency spectrum," in *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*, 2016, pp. 851–860.

[48] X. Wang, C. Yang, and S. Mao, "Phasebeat: Exploiting csi phase data for vital sign monitoring with commodity wifi devices," in *2017 IEEE 37th International Conference on Distributed Computing Systems (ICDCS)*. IEEE, 2017, pp. 1230–1239.

[49] Y. Wang, M. Huang, X. Zhu, and L. Zhao, "Attention-based lstm for aspect-level sentiment classification," in *Proceedings of the 2016 conference on empirical methods in natural language processing*, 2016, pp. 606–615.

[50] Y. Wang, K. Wu, and L. M. Ni, "Wifall: Device-free fall detection by wireless networks," *IEEE Transactions on Mobile Computing*, vol. 16, no. 2, pp. 581–594, 2016.

[51] Z. Wang, F. Xiao, N. Ye, R. Wang, and P. Yang, "A see-through-wall system for device-free human motion sensing based on battery-free rfid," *ACM Transactions on Embedded Computing Systems (TECS)*, vol. 17, no. 1, pp. 1–21, 2017.

[52] T. Wei, S. Wang, A. Zhou, and X. Zhang, "Acoustic eavesdropping through wireless vibrometry," in *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking*, 2015, pp. 130–141.

[53] T. Wei and X. Zhang, "mtrack: High-precision passive tracking using millimeter wave radios," in *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking*, 2015, pp. 117–129.

[54] G. I. Winata, O. P. Kampman, and P. Fung, "Attention-based lstm for psychological stress detection from spoken language using distant supervision," in *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2018, pp. 6204–6208.

[55] Y. Xie, J. Xiong, M. Li, and K. Jamieson, "md-track: Leveraging multi-dimensionality for passive indoor wi-fi tracking," in *The 25th Annual International Conference on Mobile Computing and Networking*, 2019, pp. 1–16.

[56] H. Xue, W. Jiang, C. Miao, F. Ma, S. Wang, Y. Yuan, S. Yao, A. Zhang, and L. Su, "Deepmv: Multi-view deep learning for device-free human activity recognition," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 4, no. 1, pp. 1–26, 2020.

[57] H. Xue, Y. Ju, C. Miao, Y. Wang, S. Wang, A. Zhang, and L. Su, "mmmesh: towards 3d real-time dynamic human mesh construction using millimeter-wave," in *Proceedings of the 19th Annual International Conference on Mobile Systems, Applications, and Services*, 2021, pp. 269–282.

[58] E. Yang, Q. He, and S. Fang, "Wink: Wireless inference of numerical keystrokes via zero-training spatiotemporal analysis," in *Proceedings of the 2022 ACM SIGSAC Conference on Computer and Communications Security*, 2022, pp. 3033–3047.

[59] L. Yang, Q. Lin, X. Li, T. Liu, and Y. Liu, "See through walls with cots rfid system!" in *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking*, 2015, pp. 487–499.

[60] Z. Yang, D. Yang, C. Dyer, X. He, A. Smola, and E. Hovy, "Hierarchical attention networks for document classification," in *Proceedings of the 2016 conference of the North American chapter of the association for computational linguistics: human language technologies*, 2016, pp. 1480–1489.

[61] S. Zhang, P. K. Sangdeh, H. Pirayesh, H. Zeng, Q. Yan, and K. Zeng, "Authiot: A transferable wireless authentication scheme for iot devices without input interface," *IEEE Internet of Things Journal*, vol. 9, no. 22, pp. 23 072–23 085, 2022.

[62] X. Zhang, Z. Li, and J. Zhang, "Synthesized millimeter-waves for human motion sensing," in *Proceedings of the Twentieth ACM Conference on Embedded Networked Sensor Systems (SenSys)*, 2022, p. 377–390.

[63] M. Zhao, T. Li, M. Abu Alsheikh, Y. Tian, H. Zhao, A. Torralba, and D. Katabi, "Through-wall human pose estimation using radio signals," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 7356–7365.

[64] M. Zhao, Y. Liu, A. Raghu, T. Li, H. Zhao, A. Torralba, and D. Katabi, "Through-wall human mesh recovery using radio signals," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2019, pp. 10 113–10 122.

[65] M. Zhao, Y. Tian, H. Zhao, M. A. Alsheikh, T. Li, R. Hristov, Z. Kabelac, D. Katabi, and A. Torralba, "Rf-based 3d skeletons," in *Proceedings of the 2018 Conference of the ACM Special Interest Group on Data Communication*, 2018, pp. 267–281.

[66] X. Zhou, X. Wan, and J. Xiao, "Attention-based lstm network for cross-lingual sentiment classification," in *Proceedings of the 2016 conference on empirical methods in natural language processing*, 2016, pp. 247–256.