SonarID: Using Sonar to Identify Fingers on a Smartwatch

Jiwan Kim
UNIST
Ulsan, Republic of Korea
kjwan4435@gmail.com

Ian Oakley
UNIST
Ulsan, Republic of Korea
ian.r.oakley@gmail.com

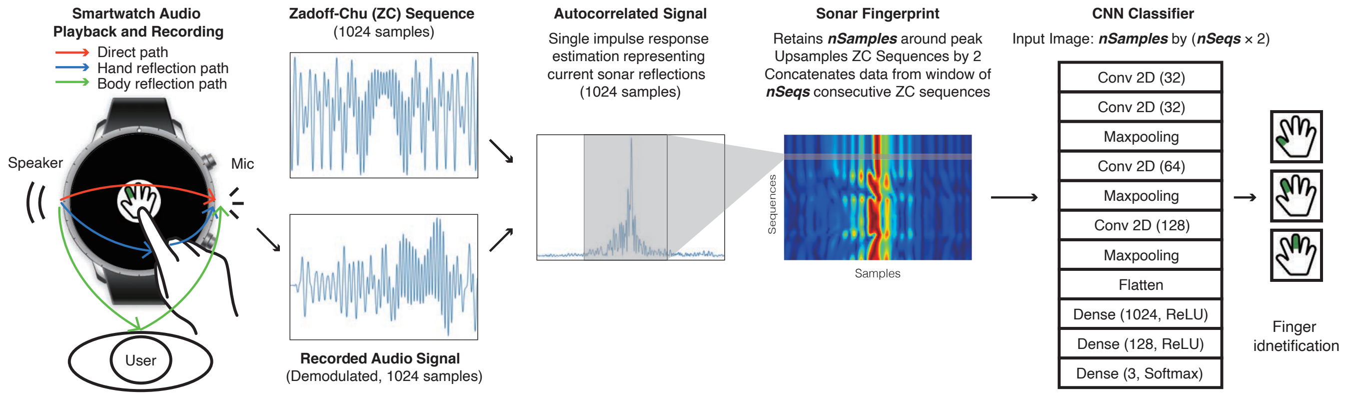

Figure 1: Overview of SonarID: during a screen touch by the thumb, index, or middle finger, a speaker on one side of a smartwatch emits an ultrasonic sonar signal (a Zadoff-Chu (ZC) sequence, modulated over a carrier wave) and a microphone on the other side receives it. The signal is demodulated and processed to create a *sonar fingerprint*: a time-varying image, composed of *nSeqs* ZC sequences, each trimmed to *nSamples* in length, of the impulse response to the signal during the touch. A deep learning model processes this data to identify which finger performed the touch.

ABSTRACT

The diminutive size of wrist wearables has prompted the design of many novel input techniques to increase expressivity. Finger identification, or assigning different functionality to different fingers, has been frequently proposed. However, while the value of the technique seems clear, its implementation remains challenging, often relying on external devices (e.g., worn magnets) or explicit instructions. Addressing these limitations, this paper explores a novel approach to natural and unencumbered finger identification on an unmodified smartwatch: sonar. To do this, we adapt an existing finger tracking smartphone sonar implementation—rather than extract finger motion, we process raw sonar fingerprints representing the complete sonar scene recorded during a touch. We capture data from 16 participants operating a smartwatch and use their sonar fingerprints to train a deep learning recognizer that identifies taps by the thumb, index, and middle fingers with an accuracy of up to 93.7%, sufficient to support meaningful application development.

CHI '22, April 29-May 5, 2022, New Orleans, LA, USA

ACM ISBN 978-1-4503-9157-3/22/04…$15.00
https://doi.org/10.1145/3491102.3501935

CCS CONCEPTS

• **Human-centered computing → Pointing**; *Sound-based input / output*; *Touch screens.*

KEYWORDS

Finger identification, smartwatch, sonar

ACM Reference Format:
Jiwan Kim and Ian Oakley. 2022. SonarID: Using Sonar to Identify Fingers on a Smartwatch. In *CHI Conference on Human Factors in Computing Systems (CHI '22), April 29-May 5, 2022, New Orleans, LA, USA.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3491102.3501935

1 INTRODUCTION

Smartwatches are increasingly powerful personal computers capable of a wide range of advanced functions such as tracking health status [22], physiological signals [10] and motor behaviors [9], mediating financial transactions [19], displaying messages and notifications [38], and supporting scheduling and navigation activities [29]. Their small size (typically in the order of 3cm by 3cm) limits the expressiveness of conventional touch screen input techniques—there is limited space to present and select on-screen targets and the fat-finger problem [32] means that much of the watch screen is obscured during interaction. To address these issues, a very wide body of research has sought to extend the input capabilities of smartwatches through techniques such as around-device interaction [2, 21], finger gesture sensing [41, 43], or augmenting the screen to detect additional touch properties such as pressure [14], contact area [27], the spatial [16] or temporal patterns of multiple touches [26], or the ability to distinguish between touching fingers [28]. In this latter approach, different functions have been assigned to each digit to support both general target selection tasks [4] and also the skilled performance of high bandwidth activities such as typing [8]. Authors have argued the modality is a good fit for the smartwatch form factor: it can be readily understood [4] and performed [8] and is rich enough to support a wide range of interaction techniques [28].

While the potential of finger identification is clear, practical, effective implementations remain challenging. In 2016, Gupta and Balakrishnan [8] argued there were no viable implementations for a smartwatch form factor and constructed their own prototype using forward-facing optical distance sensors mounted on the nails of both index and ring fingers—during a watch touch, the touching finger was identified as the one recording a nearby surface. However, while suitable for supporting empirical and design work, mounting active electronic components on the fingers is impractical and cumbersome. The situation has improved, but reliable finger identification performance is still reliant on additional worn hardware. MagTouch [28], for example, requires a user to wear a magnetic ring on their middle finger. The position of the ring is tracked by the smartwatch's magnetometer to establish the relationship between the hand and the watch and support finger identification (among index, middle, and ring fingers) during both touching and hovering events with a high level of accuracy (95.03%). However, the requirement to wear an additional device (albeit a passive ring, rather than an electronic device) in order to enable finger identification is undesirable and limits the practicality of the technique. To deal with this issue, other authors have proposed systems that do not require additional hardware. TriTap [4], for example, used the "capacitive image" [17] of each touch to infer the touching finger, a technique that achieved high levels of performance for three fingers when participants adopted specific, somewhat artificial, predefined touching poses (98%), and a reduced level of performance with natural unconstrained touches (79.4%). These articles illustrate a trade-off: the benefits of an implementation that requires no additional hardware, or places no limitations on how touches are performed, are offset by the costs of reduced finger identification accuracy.

This paper explores a novel approach to finger identification: sonar. This modality requires no additional hardware and has the ability to detect the fingers above a smartwatch, potentially supporting hover input in addition to touch. Reflecting these advantages, sonar has been previously proposed as a technique for around-device interaction on a smartwatch. FingerIO [25] presented a watch-based sonar system that tracks finger movements to support gestural or pointing input adjacent to a device. An extensive body of literature has also explored sonar on smartphones in tasks as diverse as phone grip pose identification [15], back of device input [36], breathing detection [33] and health monitoring [24]. While the smartphone literature is quite mature, the smartwatch literature is relatively limited, and we are not aware of prior work applying sonar to identify touching fingers in any context. We seek to fill this gap by presenting a sonar implementation for an off-the-shelf smartwatch and a study that captures the sonar fingerprints generated by smartwatch touches. We use these images to construct deep learning models that are able to correctly identify the touching finger with an accuracy of 93.7%.

The contributions of this paper are a sonar implementation for an off-the-shelf smartwatch, a study capturing sonar fingerprints during touch input tasks, and a description of deep learning models that are capable of using these images to reliably and accurately recognize touching fingers. These data and results indicate that sonar is a promising and effective technique for developing finger identification input systems for smartwatches that do not require modification of existing devices nor rely on users physically instrumenting their touching fingers.

2 RELATED WORK

Finger identification have been frequently proposed [30] or evaluated [6] as a mechanism for increasing the expressivity of user interaction on platforms as diverse as tabletop computers [11], smartphones [17] and smartwatches [4, 28]. The core design concept involves assigning different functions to different fingers [35, 44], or sets of fingers [6], a technique that has been demonstrated to be useful in domains as varied as text-entry (by assigning different letters to different fingers) [8] through multi-tasking (by routing input from different fingers to different applications) [7] to command specification in text editing (such as assigning different fingers to copy and paste) [5]. We argue it holds particular value for small devices such as smartwatches.

Finger identification can be implemented via numerous technologies. One approach is to enhance the touch screen, such as by making it capable of detecting fingerprints [12]. Another is to use external sensors, such as a depth camera [34], or a standard camera plus finger-worn visual markers [5, 39]. Other approaches instrument the fingers more invasively, for example, by mounting infrared [3, 23] or vibration [20] sensors on each digit. While these approaches can perform well, they are unsuitable for a real world smartwatch scenario: the small size of watch format devices precludes integration of advanced (and large) touch screen functionality, and wearing additional sensors on the body (e.g., cameras) or fingers is both impractical and undesirable. More practical approaches either use passive finger instrumentation such as a magnetic ring to track fingers [28] or derive data from the detailed analysis of the touch shapes generated by different fingers [4]. While these approaches show promise, both have limitations. Finger flexion can disrupt ring-based magnetic tracking systems [1], and effective performance using touchscreen data is reported to require the adoption of specific and somewhat artificial poses. In order to enable the full potential of the finger identification input modality on wearables, further research on the sensing techniques that can effectively enable it is currently required.

Sonar is one promising modality for this purpose. Sonar can be implemented using the speakers and microphones built in to smart devices and does not require instrumentation of the touching fingers. On mobile phones, it is well established as an input modality capable of supporting functionality as diverse as the back of device finger tracking [36], mid-air gestures [40], freehand writing [42] and breathing monitoring [33]. In order to enable this functionality, authors have typically relied on signals from multiple microphones and sought to isolate and track hand or body motion at specific

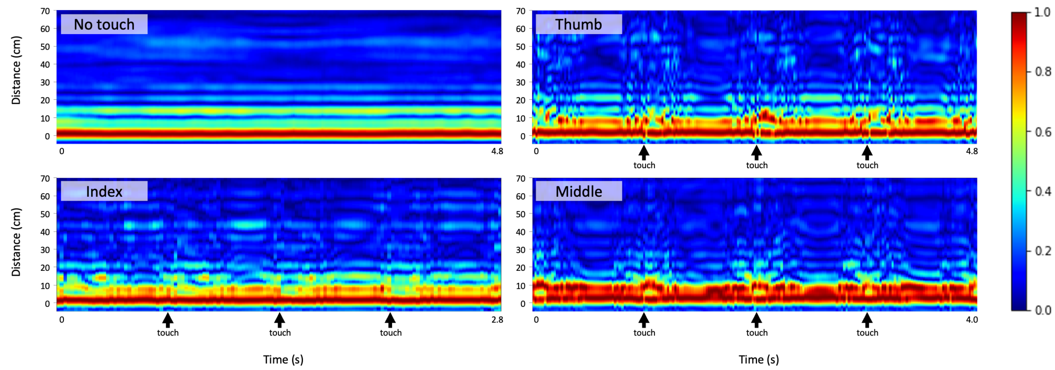

Figure 2: Examples of the sonar fingerprints, or impulse response estimations, generated during periodic smartwatch taps by each finger. Top-left shows no tapping, top-right thumb taps, bottom-left index finger taps, and bottom-right middle finger taps. Finger touches are marked under the axes on each chart. In order to facilitate visual inspection, chart x-axes have different scales (to present taps at the same spatial interval) and y-axes show only half of the auto-correlated signal.

distances and frequencies of interest using approaches such as Orthogonal Frequency Division Multiplexing (OFDM) [25], phase changes in response to multiple continuous waves [40] and impulse response estimation via auto-correlation [36]. A multiple microphone prototype in a smartwatch form factor has also shown the potential of the technique for wearables [25]. However, we know of no sonar implementations for a commercial smartwatch, none that target finger identification, and few that rely on data from a single microphone [37]—the hardware configuration available in current watches. The work in this paper seeks to address these omissions and provide insight into the potential of sonar to support finger identification using a single microphone on an off-the-shelf smartwatch.

3 SYSTEM

We implemented SonarID on an unmodified Samsung Galaxy Watch Active 2 smartwatch running Tizen. This device features a speaker on the left side, a microphone on the right side, and is capable of simultaneously playing and recording 48kHz 16 bit PCM audio. To develop our system, we selected a previously proposed sonar signal: a Zadoff-Chu (ZC) sequence (length 127, $u = 63$, up-sampled to 1024 data points) modulated into a 6kHz band over a 20.25 kHz carrier wave [36]. This signal has a number of beneficial properties. It is, by and large, inaudible. It is brief—just 21.3ms in length—and thus relatively responsive to rapid change (46.95 cycles per second). In addition, it has good auto-correlation properties (i.e., a narrow main lobe and highly attenuated side-lobes) meaning a simple auto-correlation can be used to estimate the impulse response of the signal and distinguish between peaks generated by multiple temporally proximate sonar reflections. It is reported to achieve a spatial accuracy of as low as 3.59mm in the task of tracking index finger location on the back of a smartphone [36]. This demonstrates its suitability for use in interactive systems involving close proximity between a user's hands and sonar sources and receivers. Although prior research has demonstrated the beneficial properties of this signal, we know of no work that has examined its use on a smartwatch, nor in a pose recognition task such as finger identification.

To perform finger identification, our system emits a looping ZC sequence. After a touch occurs, it segments a window of audio around that touch, demodulates the ZC signal, and cyclically estimates its impulse response by performing an auto-correlation against the original 1024 sample ZC sequence data on windows that are 512 samples (10.65ms) apart; this simple manipulation effectively up-samples the sonar measurements we take from 46.95 per second to 93.9 per second [36]. We then generate an image from multiple auto-correlation windows that depicts the changing sonar fingerprint (or impulse response pattern) during a touch. We calibrate the image to focus on the near distance by tracking the largest peak in the impulse response estimations, which is inevitably due to a combination of through-device and direct in-air audio transmission. Based on this calibration, we trim the sequence to limit our analysis to sonar reflections from predefined ranges. We refer to the two variables in this process as *nSeqs*, or the number of impulse estimations we make, and *nSamples*, or the number of samples from each impulse estimation that we retain. This former variable relates to the amount of data we analyze (e.g., the size of the temporal window used), while the latter measure corresponds to the maximum distance of the sonar reflections we consider. Figure 1 illustrates the main steps involved in this processing pipeline. Additionally, all code, scripts, and models in our system are open sources and available for download[1].

Figure 2 shows examples of data recorded using this system for no input and for a sequence of three periodic taps with the thumb, index finger, and middle finger made by a single user. Visual examination of these images suggests that the sonar fingerprints generated by these different events are sufficiently unique to support accurate finger classification. All images show a strong immediate response, representing the direct audio path. Index finger taps show limited additional reflections, while thumb and middle finger taps both present stronger proximate (5 to 10 cm) signal peaks, and distinctive sets of more distant reflections. In addition, the relatively complex, time varying nature of patterns suggests that approaches to processing the images based on identifying,

[1]https://github.com/kjwan4435/SonarID

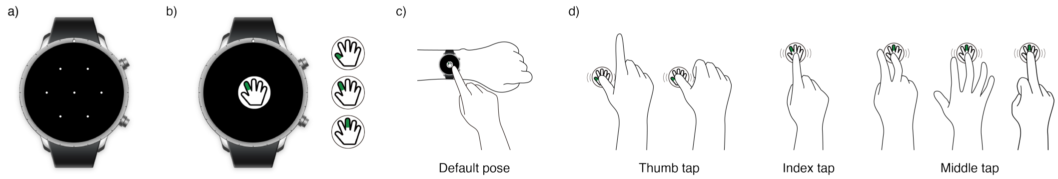

Figure 3: Study interface and interaction. It shows target positions (a), the interface during a trial (b), the index finger pose used for taps to start a trial and during the fixation period (c), and representative examples of different participants' hand posture during thumb, index and middle finger taps (d)

segmenting, and tracking key signals (e.g., the location of a single object or finger [36]) are poorly suited to the finger identification task. We suggest this is because the key differences in the signals recorded by our system represent the varying surfaces of the hand as it moves through the process of tapping the watch with different fingers. There is no single point of interest; rather, the information is contained within the changing pattern of reflections as a whole. Two key factors likely contribute to these changes: the gross movements of the hand over the watch as it brings the appropriate finger to bear, and the variations in hand pose that touches with different fingers naturally entails—retracting, extending or holding the various fingers out. Accordingly, rather than try to identify key aspects of the signal, we argue that a more profitable approach will follow recent work in room-level sonar image reconstruction [37] and apply a deep learning approach to the raw sonar data—in our case, the impulse response images. The remainder of this paper seeks to explore these assertions and determine the viability of using sonar signals to classify the finger touching a smartwatch.

4 STUDY

We conducted a study to collect sonar data from our system to support the development and evaluation of a deep learning classifier for finger identification on a smartwatch. In addition, we collected user performance data to inform a characterization of user behaviors that can shed light on the cues a classifier could use to distinguish between touches by different fingers. The study was approved by the local IRB and fully complied with both national laws and institutional regulations related to social distancing.

4.1 Participants

Sixteen participants (9 male, all right-handed, mean age of 23.94 (SD 2.32)) were recruited from the local university via online community channels. They were highly familiar with computers (4.93/5, SD 0.25) and smartphones (5/5) but relatively unfamiliar with smartwatches (1.81/5, SD: 1.05). Their hands measured 20.23cm (SD 1.42cm) in length (from base of hand to tip of middle finger), and the mean length of thumb, index, and middle fingers were, respectively, 5.81cm (SD 0.47), 7.09cm (SD 0.49) and 7.74cm (SD 0.48cm). The study took approximately an hour to complete, and each participant was compensated with approximately 13 USD in local currency.

4.2 Design

The study was designed around one key variable: fingers. We considered only three fingers, in line with prior work in this area [28]. We specifically selected the thumb, index and middle, as these digits feature in closely related research studying unencumbered finger identification on smartwatches [4] and are commonly used while interacting with pointing devices such as a mobile phone touch screen (thumb and index) or two button mouse (index and middle). In order to capture a range of touches, we laid out seven targets in a circular arrangement: one in the center, surrounded by the six others. Each target was 120 pixels (11.33 mm) in diameter, one-third of the smartwatch's screen diameter of 34mm. We indicated which finger to use for each screen touch with icons displayed directly on the target. Details of this interface are shown in Figure 3. In total, this arrangement led to 21 different trials, each with a unique finger/target combination. We arranged these trials in four blocks, each featuring a randomly ordered set composed of five repetitions of each possible trial. As such, the study captured data from 6720 trials in total: seven targets by three fingers by five repetitions by four blocks by 16 participants.

For each trial in the study, we captured the following measures: movement-time, measured from the end of the fixation period until touch down over the target; touch-time, the duration the finger was in contact with the screen; correctness, whether or not the appropriate target was selected and; raw audio. This was both emitted and captured from the start of the fixation period until 500ms after the screen was released. In addition, we recorded videos of participants' hands and watch throughout the study and asked them to self-report any erroneously completed (e.g., wrong finger) trials they noted. We did not incorporate any further independent measures of the correctness of tapping fingers. Following prior work in this area [4], we relied on the simplicity of the study task to ensure that the vast majority of the study trials were completed correctly.

4.3 Procedure

The experiment took place in an empty classroom with participants seated in front of an empty desk. To prevent fatigue, they rested their wrist on the desk. Participants first read the instructions and signed consent. They then donned the smartwatch on their left wrist and had a maximum of five minutes to practice the three finger touches (thumb, index, middle). They were instructed to determine the most comfortable and effective input actions for completing the study tasks—this was important as many participants were unfamiliar with both smartwatches and the use of either thumb or middle finger to perform taps. This practice stage helped reduce variability in the study tasks as it allowed participants to experiment with different approaches to making the finger taps. However, we note that we did not restrict participants' hand, arm, or finger

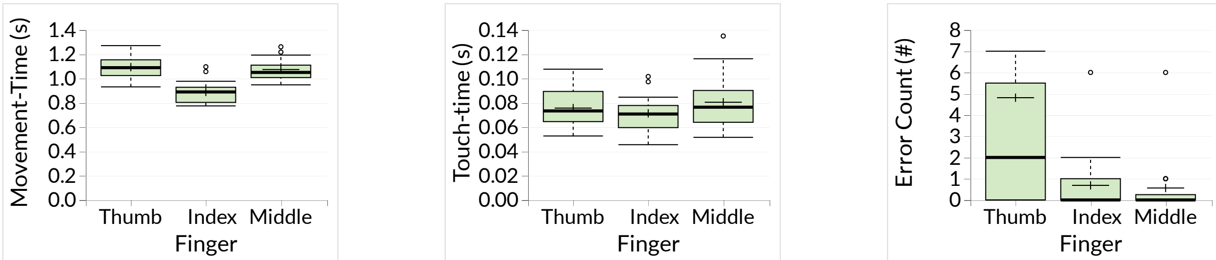

Figure 4: Study data showing movement-times (left), touch-times (center), and error counts (right) for touches with the thumb, index, and middle fingers.

poses—they were free to tap the watch any way they felt was comfortable. After indicating they were ready, the study began. Each trial involved tapping the screen with the index finger, then hovering over the watch for a fixation period (500ms), after which time the trial target and finger were shown. The participant then tapped the target with the appropriate finger, and the next trial began. Incomplete trials timed out after three seconds. Participants were able to rest at any time between trials, and there were three enforced breaks of two minutes, one between each of the four trial blocks. At the end of the study they completed demographics.

4.4 Results

We recorded a total of 6530 (97.17%) correctly completed trials and 97 errors (1.44%) in which the on-screen target was not correctly tapped. In addition, we logged 30 errors (0.45%) in which participants self-reported erroneously tapping the screen (e.g., with the wrong finger) and 63 errors (0.94%) in which technical problems led to data loss. We opted not to analyze self-reported errors due to their sparsity and technical errors due to their lack of relevance to user performance. Data for movement-time and touch-time from correctly completed trials, and trial error count, are shown in Figure 4. Movement- and touch-time data were normally distributed and upheld sphericity assumptions, so we analyzed them with one-way repeated measures ANOVA on the finger variable. Movement-time led to a significant main effect (F (2, 15) = 123.578, p<0.001, $\hat{\eta}^2_G$=0.493), while touch-time did not (F (2, 15) = 2.715, p=0.082, $\hat{\eta}^2_G$=0.043). Post-hoc t-tests on movement-time, incorporating Bonferroni corrections, indicated that taps with the index finger were faster than with the thumb or middle finger (both p<0.001). This suggests that thumb and middle finger taps in our task both involved additional hand movements prior to contact with the screen and that these motions exacted a modest, but stable, time cost of approximately 200ms. These motions are likely due to the fact that the index finger was mandated for the taps used to start each trial. As prior work has indicated sonar sensing systems on smart devices are highly sensitive to hand and finger motion [36], we note these variations will also likely constitute a key feature that a classifier can use to distinguish between taps by different fingers, with thumb taps involving a descending rightward motion of the hand, index taps a simple downward motion and middle taps a descending leftward motion.

We then examined errors. Error counts were not normally distributed; in fact, three participants contributed 51 (66%) of the errors, with numerous others achieving perfect or near-perfect performance. Additionally, the vast majority of errors (92%) occurred on the middle row of targets—on the leftmost (23%), center (54%), or rightmost (15%) targets. Reflecting the unevenness of this data, we analyzed error counts using a Friedman test on the finger variable. It revealed significant differences ($\chi^2(2)$ = 5.848, p=0.008), which follow up Wilcoxon tests (applying an α level of 0.0167 to emulate Bonferroni correction) indicated were due to the index (Z=2.5, p=0.007) and middle (Z=4.0, p=0.011) finger leading to fewer wrong target selections than the thumb. We conclude that although error rates in the study were generally low (at 1.44% in total), reflecting the simplicity of the study task, there was a subset of participants who experienced some degree of challenge in the task of accurately thumb tapping buttons located in the center row of the watch screen. To support optimal performance for all users, targets intended for thumb tapping might therefore be better located at the top or bottom of a smartwatch screen. We could not infer potential features that might support finger classification from this data due, in part, to its sparsity.

Finally, we took notes live and examined the study video recordings to informally catalog variations in participants' performance of the three taps in the study. Index finger taps were uniformly performed simply with a single, isolated, outstretched digit. For thumb taps, performance was also quite uniform, with 14 participants opting to make a "V pose" with the thumb and index, with other fingers tucked into a loose fist. The remaining two participants (P10, P11) performed thumb taps after first retracting their index finger to join the others in their fists. There was somewhat more diversity in middle finger taps, with 11 participants making a "V pose" with index and middle, two participants extending the middle finger alone (again, P10 and P11), and the remaining three (P2, P5, P12) extending all fingers, loosely splaying their hand, during a middle finger tap. These poses are illustrated in Figure 3. While we expected the diversity of these user-selected poses to present challenges to our goal of developing a sonar based recognizer for finger identification, the similarities among many participants' input may also lead to generation of consistent classes of sonar reflection. For most participants, thumb touches meant their hand was situated to the right of the watch, while index touches involved a single digit above the watch and middle touches led to the fingers or hand being on the left of (or covering) the watch. We suggest the patterns of sonar reflections generated by these different poses may be sufficiently distinct to support reliable finger classification.

Table 1: Accuracy (in %) for different sonar ranges, expressed in *nSamples*, the length of the ZC sequence used. *nSeqs* is set to 60/40.

Signal Range	Near		Mid		Far
nSamples (#)	25	50	100	200	300
Distance (cm)	4.5	9	18	36	54
Accuracy (%)	86.06	89.43	91.73	93.26	93.72

4.5 Preprocessing

We processed the recorded sound for each trial according to the procedures set out in Section 3. In brief, the recorded ZC sequence data was demodulated from the carrier wave, then cyclically auto-correlated with the original 1024 sample ZC sequence with a hop size of 512. Each of these auto-correlations yielded an estimate of the impulse response to the signal that encompasses reflections from a maximum distance of approximately 368cm [36]. We concatenated a sequence of these estimations to form a sonar fingerprint—a time-varying image representing the sonar reflections recorded throughout a touch. We defined two variables in this process: *nSamples*, referring to the number of samples used from each impulse response estimation, and *nSeqs*, referring to the number of estimations we concatenate. These correspond to, respectively, the physical range at which we capture sonar reflections and the period of time in which we capture them. We explored the impact of varying the value of these parameters in a grid search procedure (see Tables 1 and 2). In addition, Appendix A depicts representative examples of sonar fingerprints captured from different users for each of the three finger touches.

4.6 Classifier and Classification Performance

We sought to identify the fingers involved in screen touches from our sonar fingerprints using a 2D convolution neural network (ConvNet) [18]. Due to the similarity between our data and traditional sonograms, we based our system on an existing design for detecting visual patterns in audio spectrograms [31]. To support feature learning, we selected a model composed of three convolution blocks, with each block containing a convolution, Rectified Linear unit (ReLU) activation, and max-pooling layer. We consecutively increased the number of filters (from 32 to 64 then 128) and used a 3x3 kernel in the convolution layers, while we used a 2x2 kernel and a stride length of 2 in the max-pooling layers. After we flattened the output from the last convolution block, we included two fully connected hidden layers with, respectively, 1024 and 128 units. A final fully connected layer performed the multi-classification into thumb, index, or middle finger classes. Figure 1 shows this structure.

We first constructed general models using data from all participants to explore the *nSamples* and *nSeqs* parameters via a grid search procedure. In this process, we used 80% of the data for training and reserved the remaining 20% for final testing. Using the training data set, we conducted five-fold cross validation procedures. We selected values for *nSamples* to be between 25 to 300, encompassing reflections from between 4.5 cm (single reflections from the finger) to 54 cm (multiple reflections from the body, arm, hand and fingers). We considered *nSeqs* values of between 20 and 100, specifying periods with a duration of between 213ms and 1065ms. To shed light on which periods contain the most salient information, we considered intervals both before the temporal midpoint of a touch (213-639ms), after the temporal midpoint of a touch (213-426ms), and the combination of these ranges. We omitted the 639ms period after the touch mid-point due to the high latency it entails and the fact that we terminated audio capture 500ms after finger up. The performance of models constructed during this grid search procedure on our final test data set are shown in Tables 1 and 2. Perhaps unsurprisingly, they indicate that peak accuracy (93.7%) is achieved with the most data: the configuration that includes the greatest range (*nSamples* set to 300, 54cm reflections) and time period (*nSeqs* set to 100, including 60 samples before and 40 samples after touch midpoint, 1065ms). However, we note that performance with lower distances remains high—data from a range of just 9cm, encompassing the finger and hand, achieves an accuracy of 89.43%. We suggest the benefits from including further ranges may be due, in part, the presence of echoes or multiple reflections from the hand and wrist. In terms of time, periods before the touch midpoint showed modest improvements over those after it, while the combination of both periods led to peaks; we conclude that motions of both finger approach and retraction from the screen contained valuable information to support classification. In addition, the short periods around the touch midpoint, where finger and hand motions are likely small or slow, showed low performance. The key sonar features therefore likely relate to the changing sonar reflections that are recorded as the hand moves. Data from more static poses were less salient. Based on these results, we selected the optimally performing general model configuration, using data from a range of 54cm (*nSamples* = 300) captured over periods of 1065ms (*nSeqs* = 100), for all further tests.

To validate performance, we produced two further sets of models: individual models and Leave One Out Cross Validation (LOOCV) models. In the individual models, each participant's own data was used to produce a model. We followed the same procedures used to create the general model: an 8:2 train/test data split and five-fold cross-validation procedures on the training set. For the LOOCV models, data from each participant served as a final test set for a model trained using five-fold cross validation on data from all other participants (essentially, a 15:1 train/test split). Data from these models are shown in Table 3. Individual models showed a similar performance profile to the general model: mean accuracy was also 93.7%. We suggest that individual models were unable to improve on the general model performance due to the relatively sparsity of data in each. Data from the LOOCV models reinforces this conclusion. While fair (85.37%), it is notably reduced from that achieved using the general or per-user models. This suggests that, within our relatively small sample, the sonar fingerprints created by each users' tapping behavior were somewhat specific to that user. In order to achieve LOOCV performance equivalent to that in the general model, or individual model performance that exceeds it, we would likely need to sample more data from more users.

It is also worth reflecting on the distribution of classification errors among the finger classes. Figure 5 shows this for the general model and the mean performance for the sets of both individual and LOOCV models. Similar trends can be observed. Classification performance peaks with the index finger while thumb and middle are more frequently confused. This effect is particularly prominent

Table 2: Accuracy (in %) for different data capture periods, expressed in terms of both *nSeqs*, the number of impulse estimations performed, and time (ms). In these results, *nSamples* is set to 300.

Signal Period	**Before Touch**			**After Touch**		**Before/After Touch**		
nSeqs (#)	20	40	60	20	40	20/20	40/40	60/40
Time (ms)	213	426	639	213	426	213/213	426/426	639/426
Accuracy (%)	62.94	86.37	89.59	60.95	82.39	69.98	90.81	93.72

Table 3: Accuracy (in %) for individual and LOOCV models, including mean (μ) and standard deviation (σ) from all participants.

	P1	**P2**	**P3**	**P4**	**P5**	**P6**	**P7**	**P8**	**P9**	**P10**	**P11**	**P12**	**P13**	**P14**	**P15**	**P16**	μ	σ
Individual (%)	92.32	92.49	93.2	90.93	95.32	94.04	92.48	95.94	93.27	95.64	93.44	94.86	89.07	96.62	95.04	94.74	93.71	1.98
LOOCV (%)	88.25	74.51	89.78	78.28	85.47	81.62	88.62	95.23	85.54	89.1	91.99	86.31	79.64	80.47	83.37	87.72	85.37	5.43

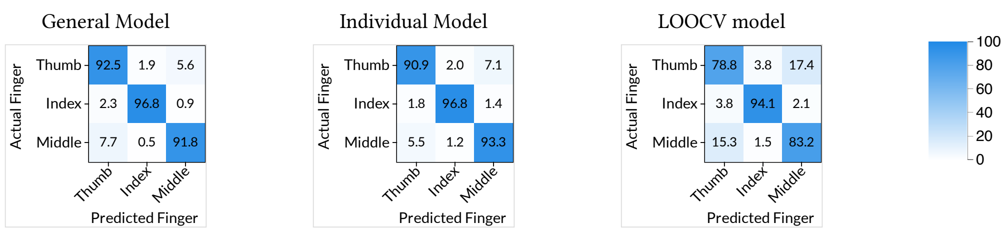

Figure 5: Confusion matrices for SonarID classifiers (% accuracy). Left: general model; center: mean individual model; right: mean LOOCV model.

in the LOOCV models, where thumb and middle fingers are correctly recognized just 78.8% and 83.2% of the time. We can extract a number of implications from this result. First and foremost, it suggests that performance in a two class task (as used by Gupta and Balakrishnan [8]) of separating the index taps from those of other fingers would be high: up to 96% even in the LOOCV models. Secondly, it likely reflects the diversity of strategies we observed for middle and thumb taps—while index finger taps were universal, there were several different high level strategies (such as tucking in or extending unused fingers) during thumb and middle taps. This diversity may have increased the difficulty of the classification task. As such, larger training sets may be able to more clearly separate finger touches based on various different strategies. Alternatively, customized user models may be able to more accurately cater to a given user's particular touching style.

5 DISCUSSION AND CONCLUSION

This paper presents a system that recognizes the finger tapping an off-the-shelf smartwatch using a sonar scene generated and sensed via the device's built-in speaker and microphone. It achieves a peak accuracy of 93.7% for both individual and general models and a LOOCV accuracy of 85.4%. These results compare favourably to prior work implementing finger identification using standard smartwatches. TriTap [4], for example, uses touchscreen data for a similar classification task and achieves an accuracy of 79.4% using individual models and natural touches. Furthermore, the performance we report is only marginally reduced compared to systems that track worn objects, such as the magnetic ring in MagTouch [28] that supports a recognition rate (among index, middle, and ring fingers) of 95.03%. We argue that the convenience and practicality of enabling input with an unencumbered touching hand offer advantages over systems that require users to wear additional hardware.

It is also worth discussing the detailed performance profile of our results. We show strong performance in classifying the index finger and difficulty distinguishing between thumb and middle. TriTap's [4] capacitive screen implementation shows the strongest performance in distinguishing the thumb and confuses the index and middle. A multi-modal combination of these two approaches would likely be highly complementary. Similarly, SonarID's ability to classify thumb taps may be able to extend MagTouch's [28] finger-only approach with an additional digit. We see strong potential in combining these modalities in the future. Data from the grid search over temporal periods for capturing a sonar signal (*nSeqs*) also presents implications for the design of interactive systems. It suggests that peak performance will require continual emission and recording of sound—the highest performance we observed uses data from both before and after touch. Although prior work has argued that power consumption for sonar systems is reasonable on a smartphone [36], the impact on smartwatch battery life may be more extreme—during the intensive watch use in our hour long study, watch battery level declined by 8%-10% for each participant. A more efficient approach of recording signals only after a touch offers reduced classification accuracy (of 82.39%), suggesting a potential trade-off between power consumption and classification performance. We also note that using sonar data from after a touch implies a latency (roughly equivalent, in our studies, to the 500ms commonly used for long tap) before events can be detected. While that latency can be avoided by relying on data gathered up until a touch occurs, this approach also leads to a more modest reduction in classification accuracy (to 89.59%). Future studies and system designs will need to flesh out and balance these concerns.

There are a number of further limitations to this work; these signpost next steps for this project. The simplest relates to the abundancy of the data we rely on—results from both our individual and LOOCV models suggest that improved performance could be achieved with larger training sets. While this could be achieved simply via extended empirical work, we also highlight the need to explore data augmentation techniques on our existing data set [31]. Such techniques may enable us to improve the performance of our classifiers without requiring further studies. Similarly, we may be able to improve or optimize the classifier itself. One key way which this could be done is by reducing the number of parameters in the models—this is currently approximately 58M, a figure reported to be high for use on a wearable device [13]. To explore the viability of lowering this figure, we constructed two new versions of our general model. In one we removed the first dense network layer (1024 units, reducing the total parameters to 7.3M) and in the second we also added a new max-pooling layer after the initial Conv2D layer (leading to 1.7M parameters). Accuracy in these models remained high at 93.5% and 92.2%, respectively. These results highlight both the robustness of our solution and the potential for optimizing performance further by exploring more sophisticated classifier architectures.

While these technical activities may improve and further validate our system, we also note further studies are inevitable. Although much work in finger identification uses a single study [4] or pose [8], other authors note the advantages in terms of robustness and validity that can be realized by sampling data from various situations [28]. While the seated, hand-ready pose we use in this work is both common and representative, a clear next step for this work is to capture data from more diverse situations and environments. These should include while standing, with the arms in various poses, and with an increased diversity of tapping styles, such as fully separated single touches that each involve a finger approaching the watch independently. While performance may change if data is captured from more diverse poses, we note that our current results indicate that proximate sonar reflections from 4.5cm and 9cm, involving just the touching fingers and hand (see Table 1), lead to relatively high accuracy levels of up to 89.43%. This result suggests SonarID may be resilient to changes in upper arm or body pose as such variations would likely impact only more distant sonar reflections. In addition, it will be important to explore the robustness of our technique to various forms of environmental disturbance such as as different ambient noise conditions (situations in which which closely related prior sonar systems have performed well [36]), or wind—a form of environmental disturbance widely acknowledged to reduce the signal to noise ratio of sonar systems. Our future plans for this project involve addressing these issues. By capturing more data in diverse settings and exploring techniques to augment that data, we believe we can construct classifiers with improved accuracy and increased validity. Doing so will ensure the technique we describe works not only on off-the-shelf devices but also in real-world settings.

ACKNOWLEDGMENTS

Thanks to all our study participants. This paper was partly supported by the Bio & Medical Technology Development Program of the National Research Foundation (NRF) and funded by the Korean government (MSIT) (NRF-2021M3A9E4080780) and a Korea Institute for Advancement of Technology (KIAT) grant funded by the Korean Government (MOTIE) (P0012725, The Competency Development Program for Industry Specialist).

REFERENCES

[1] Daniel Ashbrook, Patrick Baudisch, and Sean White. 2011. *Nenya: Subtle and Eyes-Free Mobile Input with a Magnetically-Tracked Finger Ring.* Association for Computing Machinery, New York, NY, USA, 2043–2046. https://doi.org/10.1145/1978942.1979238

[2] Alex Butler, Shahram Izadi, and Steve Hodges. 2008. SideSight: Multi-"touch" Interaction around Small Devices. In *Proceedings of the 21st Annual ACM Symposium on User Interface Software and Technology* (Monterey, CA, USA) *(UIST '08)*. Association for Computing Machinery, New York, NY, USA, 201–204. https://doi.org/10.1145/1449715.1449746

[3] Philipp Ewerling, Alexander Kulik, and Bernd Froehlich. 2012. Finger and Hand Detection for Multi-Touch Interfaces Based on Maximally Stable Extremal Regions. In *Proceedings of the 2012 ACM International Conference on Interactive Tabletops and Surfaces* (Cambridge, Massachusetts, USA) *(ITS '12)*. Association for Computing Machinery, New York, NY, USA, 173–182. https://doi.org/10.1145/2396636.2396663

[4] Hyunjae Gil, DoYoung Lee, Seunggyu Im, and Ian Oakley. 2017. *TriTap: Identifying Finger Touches on Smartwatches.* Association for Computing Machinery, New York, NY, USA, 3879–3890. https://doi.org/10.1145/3025453.3025561

[5] Alix Goguey, Géry Casiez, Daniel Vogel, Fanny Chevalier, Thomas Pietrzak, and Nicolas Roussel. 2014. A Three-Step Interaction Pattern for Improving Discoverability in Finger Identification Techniques. In *Proceedings of the Adjunct Publication of the 27th Annual ACM Symposium on User Interface Software and Technology* (Honolulu, Hawaii, USA) *(UIST'14 Adjunct)*. Association for Computing Machinery, New York, NY, USA, 33–34. https://doi.org/10.1145/2658779.2659100

[6] Alix Goguey, Mathieu Nancel, Géry Casiez, and Daniel Vogel. 2016. *The Performance and Preference of Different Fingers and Chords for Pointing, Dragging, and Object Transformation.* Association for Computing Machinery, New York, NY, USA, 4250–4261. https://doi.org/10.1145/2858036.2858194

[7] Aakar Gupta, Muhammed Anwar, and Ravin Balakrishnan. 2016. Porous Interfaces for Small Screen Multitasking Using Finger Identification. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology* (Tokyo, Japan) *(UIST '16)*. Association for Computing Machinery, New York, NY, USA, 145–156. https://doi.org/10.1145/2984511.2984557

[8] Aakar Gupta and Ravin Balakrishnan. 2016. *DualKey: Miniature Screen Text Entry via Finger Identification.* Association for Computing Machinery, New York, NY, USA, 59–70. https://doi.org/10.1145/2858036.2858052

[9] Yuvin Ha, Maria Karyda, and Andrés Lucero. 2020. *Exploring Virtual Rewards in Real Life: A Gimmick or a Motivational Tool for Promoting Physical Activity?* Association for Computing Machinery, New York, NY, USA, 1847–1858. https://doi.org/10.1145/3357236.3395477

[10] Tian Hao, Chongguang Bi, Guoliang Xing, Roxane Chan, and Linlin Tu. 2017. MindfulWatch: A Smartwatch-Based System For Real-Time Respiration Monitoring During Meditation. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 1, 3, Article 57 (Sept. 2017), 19 pages. https://doi.org/10.1145/3130922

[11] Chris Harrison, Julia Schwarz, and Scott E. Hudson. 2011. TapSense: Enhancing Finger Interaction on Touch Surfaces. In *Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology* (Santa Barbara, California, USA) *(UIST '11)*. Association for Computing Machinery, New York, NY, USA, 627–636. https://doi.org/10.1145/2047196.2047279

[12] Christian Holz and Patrick Baudisch. 2013. Fiberio: A Touchscreen That Senses Fingerprints. In *Proceedings of the 26th Annual ACM Symposium on User Interface Software and Technology* (St. Andrews, Scotland, United Kingdom) *(UIST '13)*. Association for Computing Machinery, New York, NY, USA, 41–50. https://doi.org/10.1145/2501988.2502021

[13] Andrew Howard, Mark Sandler, Grace Chu, Liang-Chieh Chen, Bo Chen, Mingxing Tan, Weijun Wang, Yukun Zhu, Ruoming Pang, Vijay Vasudevan, Quoc V. Le, and Hartwig Adam. 2019. Searching for MobileNetV3. arXiv:1905.02244 [cs.CV]

[14] Min-Chieh Hsiu, Da-Yuan Huang, Chi An Chen, Yu-Chih Lin, Yi-ping Hung, De-Nian Yang, and Mike Chen. 2016. ForceBoard: Using Force as Input Technique on Size-Limited Soft Keyboard. In *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct* (Florence, Italy) *(MobileHCI '16)*. Association for Computing Machinery, New York, NY, USA, 599–604. https://doi.org/10.1145/2957265.2961827

[15] Namhyun Kim, Junseong Lee, Joyce Jiyoung Whang, and Jinkyu Lee. 2020. SmartGrip: grip sensing system for commodity mobile devices through sound signals. *Personal and Ubiquitous Computing* 24, 5 (2020), 643–654.
[16] Benjamin Lafreniere, Carl Gutwin, Andy Cockburn, and Tovi Grossman. 2016. *Faster Command Selection on Touchscreen Watches*. Association for Computing Machinery, New York, NY, USA, 4663–4674. https://doi.org/10.1145/2858036.2858166
[17] Huy Viet Le, Sven Mayer, and Niels Henze. 2019. Investigating the Feasibility of Finger Identification on Capacitive Touchscreens Using Deep Learning. In *Proceedings of the 24th International Conference on Intelligent User Interfaces* (Marina del Ray, California) *(IUI '19)*. Association for Computing Machinery, New York, NY, USA, 637–649. https://doi.org/10.1145/3301275.3302295
[18] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner. 1998. Gradient-based learning applied to document recognition. *Proc. IEEE* 86, 11 (1998), 2278–2324. https://doi.org/10.1109/5.726791
[19] Chris Xiaoxuan Lu, Bowen Du, Hongkai Wen, Sen Wang, Andrew Markham, Ivan Martinovic, Yiran Shen, and Niki Trigoni. 2018. Snoopy: Sniffing Your Smartwatch Passwords via Deep Sequence Learning. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 1, 4, Article 152 (Jan. 2018), 29 pages. https://doi.org/10.1145/3161196
[20] Damien Masson, Alix Goguey, Sylvain Malacria, and Géry Casiez. 2017. WhichFingers: Identifying Fingers on Touch Surfaces and Keyboards Using Vibration Sensors. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology* (Québec City, QC, Canada) *(UIST '17)*. Association for Computing Machinery, New York, NY, USA, 41–48. https://doi.org/10.1145/3126594.3126619
[21] Jess McIntosh, Paul Strohmeier, Jarrod Knibbe, Sebastian Boring, and Kasper Hornbæk. 2019. *Magnetips: Combining Fingertip Tracking and Haptic Feedback for Around-Device Interaction*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3290605.3300638
[22] Tejaswini Mishra, Meng Wang, Ahmed A Metwally, Gireesh K Bogu, Andrew W Brooks, Amir Bahmani, Arash Alavi, Alessandra Celli, Emily Higgs, Orit Dagan-Rosenfeld, et al. 2020. Pre-symptomatic detection of COVID-19 from smartwatch data. *Nature biomedical engineering* 4, 12 (2020), 1208–1220.
[23] Sundar Murugappan, Vinayak, Niklas Elmqvist, and Karthik Ramani. 2012. Extended Multitouch: Recovering Touch Posture and Differentiating Users Using a Depth Camera. In *Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology* (Cambridge, Massachusetts, USA) *(UIST '12)*. Association for Computing Machinery, New York, NY, USA, 487–496. https://doi.org/10.1145/2380116.2380177
[24] Rajalakshmi Nandakumar, Shyamnath Gollakota, and Jacob E Sunshine. 2019. Opioid overdose detection using smartphones. *Science translational medicine* 11, 474 (2019), 10 pages.
[25] Rajalakshmi Nandakumar, Vikram Iyer, Desney Tan, and Shyamnath Gollakota. 2016. *FingerIO: Using Active Sonar for Fine-Grained Finger Tracking*. Association for Computing Machinery, New York, NY, USA, 1515–1525. https://doi.org/10.1145/2858036.2858580
[26] Ian Oakley, DoYoung Lee, MD. Rasel Islam, and Augusto Esteves. 2015. *Beats: Tapping Gestures for Smart Watches*. Association for Computing Machinery, New York, NY, USA, 1237–1246. https://doi.org/10.1145/2702123.2702226
[27] Ian Oakley, Carina Lindahl, Khanh Le, DoYoung Lee, and MD. Rasel Islam. 2016. *The Flat Finger: Exploring Area Touches on Smartwatches*. Association for Computing Machinery, New York, NY, USA, 4238–4249. https://doi.org/10.1145/2858036.2858179
[28] Keunwoo Park, Daehwa Kim, Seongkook Heo, and Geehyuk Lee. 2020. *MagTouch: Robust Finger Identification for a Smartwatch Using a Magnet Ring and a Built-in Magnetometer*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3313831.3376234
[29] Martin Perebner, Haosheng Huang, and Georg Gartner. 2019. Applying user-centred design for smartwatch-based pedestrian navigation system. *Journal of location based services* 13, 3 (2019), 213–237.
[30] Quentin Roy, Yves Guiard, Gilles Bailly, Éric Lecolinet, and Olivier Rioul. 2015. Glass+Skin: An Empirical Evaluation of the Added Value of Finger Identification to Basic Single-Touch Interaction on Touch Screens. In *Human-Computer Interaction – INTERACT 2015*, Julio Abascal, Simone Barbosa, Mirko Fetter, Tom Gross, Philippe Palanque, and Marco Winckler (Eds.). Springer International Publishing, Cham, 55–71.
[31] Justin Salamon and Juan Pablo Bello. 2017. Deep Convolutional Neural Networks and Data Augmentation for Environmental Sound Classification. *IEEE Signal Processing Letters* 24, 3 (2017), 279–283. https://doi.org/10.1109/LSP.2017.2657381
[32] Katie A. Siek, Yvonne Rogers, and Kay H. Connelly. 2005. Fat Finger Worries: How Older and Younger Users Physically Interact with PDAs. In *Human-Computer Interaction - INTERACT 2005*, Maria Francesca Costabile and Fabio Paternò (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 267–280.
[33] Xingzhe Song, Boyuan Yang, Ge Yang, Ruirong Chen, Erick Forno, Wei Chen, and Wei Gao. 2020. SpiroSonic: Monitoring Human Lung Function via Acoustic Sensing on Commodity Smartphones. In *Proceedings of the 26th Annual International Conference on Mobile Computing and Networking* (London, United Kingdom) *(MobiCom '20)*. Association for Computing Machinery, New York, NY, USA, Article 52, 14 pages. https://doi.org/10.1145/3372224.3419209
[34] Srinath Sridhar, Anders Markussen, Antti Oulasvirta, Christian Theobalt, and Sebastian Boring. 2017. *WatchSense: On- and Above-Skin Input Sensing through a Wearable Depth Sensor*. Association for Computing Machinery, New York, NY, USA, 3891–3902. https://doi.org/10.1145/3025453.3026005
[35] Atsushi Sugiura and Yoshiyuki Koseki. 1998. A User Interface Using Fingerprint Recognition: Holding Commands and Data Objects on Fingers. In *Proceedings of the 11th Annual ACM Symposium on User Interface Software and Technology* (San Francisco, California, USA) *(UIST '98)*. Association for Computing Machinery, New York, NY, USA, 71–79. https://doi.org/10.1145/288392.288575
[36] Ke Sun, Ting Zhao, Wei Wang, and Lei Xie. 2018. VSkin: Sensing Touch Gestures on Surfaces of Mobile Devices Using Acoustic Signals. In *Proceedings of the 24th Annual International Conference on Mobile Computing and Networking* (New Delhi, India) *(MobiCom '18)*. Association for Computing Machinery, New York, NY, USA, 591–605. https://doi.org/10.1145/3241539.3241568
[37] Alex Turpin, Valentin Kapitany, Jack Radford, Davide Rovelli, Kevin Mitchell, Ashley Lyons, Ilya Starshynov, and Daniele Faccio. 2021. 3D Imaging from Multipath Temporal Echoes. *Physical Review Letters* 126, 17 (Apr 2021), 6 pages. https://doi.org/10.1103/physrevlett.126.174301
[38] Aku Visuri, Niels van Berkel, Jorge Goncalves, Reza Rawassizadeh, Denzil Ferreira, and Vassilis Kostakos. 2021. Understanding usage style transformation during long-term smartwatch use. *Personal and Ubiquitous Computing* 25 (2021), 535–549.
[39] Jingtao Wang and John Canny. 2004. FingerSense: Augmenting Expressiveness to Physical Pushing Button by Fingertip Identification. In *CHI '04 Extended Abstracts on Human Factors in Computing Systems* (Vienna, Austria) *(CHI EA '04)*. Association for Computing Machinery, New York, NY, USA, 1267–1270. https://doi.org/10.1145/985921.986040
[40] Wei Wang, Alex X. Liu, and Ke Sun. 2016. Device-Free Gesture Tracking Using Acoustic Signals. In *Proceedings of the 22nd Annual International Conference on Mobile Computing and Networking* (New York City, New York) *(MobiCom '16)*. Association for Computing Machinery, New York, NY, USA, 82–94. https://doi.org/10.1145/2973750.2973764
[41] Hongyi Wen, Julian Ramos Rojas, and Anind K. Dey. 2016. *Serendipity: Finger Gesture Recognition Using an Off-the-Shelf Smartwatch*. Association for Computing Machinery, New York, NY, USA, 3847–3851. https://doi.org/10.1145/2858036.2858466
[42] Kaishun Wu, Qiang Yang, Baojie Yuan, Yongpan Zou, Rukhsana Ruby, and Mo Li. 2021. EchoWrite: An Acoustic-Based Finger Input System Without Training. *IEEE Transactions on Mobile Computing* 20, 5 (2021), 1789–1803. https://doi.org/10.1109/TMC.2020.2973094
[43] Yu Zhang, Tao Gu, Chu Luo, Vassilis Kostakos, and Aruna Seneviratne. 2018. FinDroidHR: Smartwatch Gesture Input with Optical Heartrate Monitor. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 2, 1, Article 56 (March 2018), 42 pages. https://doi.org/10.1145/3191788
[44] Jingjie Zheng and Daniel Vogel. 2016. *Finger-Aware Shortcuts*. Association for Computing Machinery, New York, NY, USA, 4274–4285. https://doi.org/10.1145/2858036.2858355

A SONAR FINGERPRINT SAMPLES

Figure 6 shows a set of sonar fingerprints captured from different users in our study. We include four correct classifications for each of thumb, index and middle finger: the first four rows of images in each column. The first three of these rows show relatively similar images, while those on the fourth row are somewhat distinct. They come from a single participant (P11) who adopted a less common tapping style in which each touching finger was extended alone, with all other fingers tucked into the palm (see Figure 3). Additionally, the final two rows of images show examples of all possible finger misclassifications. These images suggest that taps by individual fingers, such as the index finger for the majority of participants and all fingers for P11 (fourth row), lead to a simple, proximate set of sonar reflections: a single extended finger presents a limited and nearby surface for sound to strike. Thumb and middle taps, in which other fingers are extended for most users, present a larger set of surfaces and more complex reflections, with middle finger taps leading to noticeably more distant reflections than thumb taps. This suggests that aspects of hand pose are reliably captured in our sonar fingerprints. Additionally, our study involved a task in which an index finger tap was used to start a trial. Cues related to subsequent movements of the hand rightward or leftward to position, respectively, the thumb or middle finger over the screen prior to tapping likely lead to the differences, and reliable classification performance, observed in cases where the different fingers were presented alone (e.g., P11, fourth row of images).

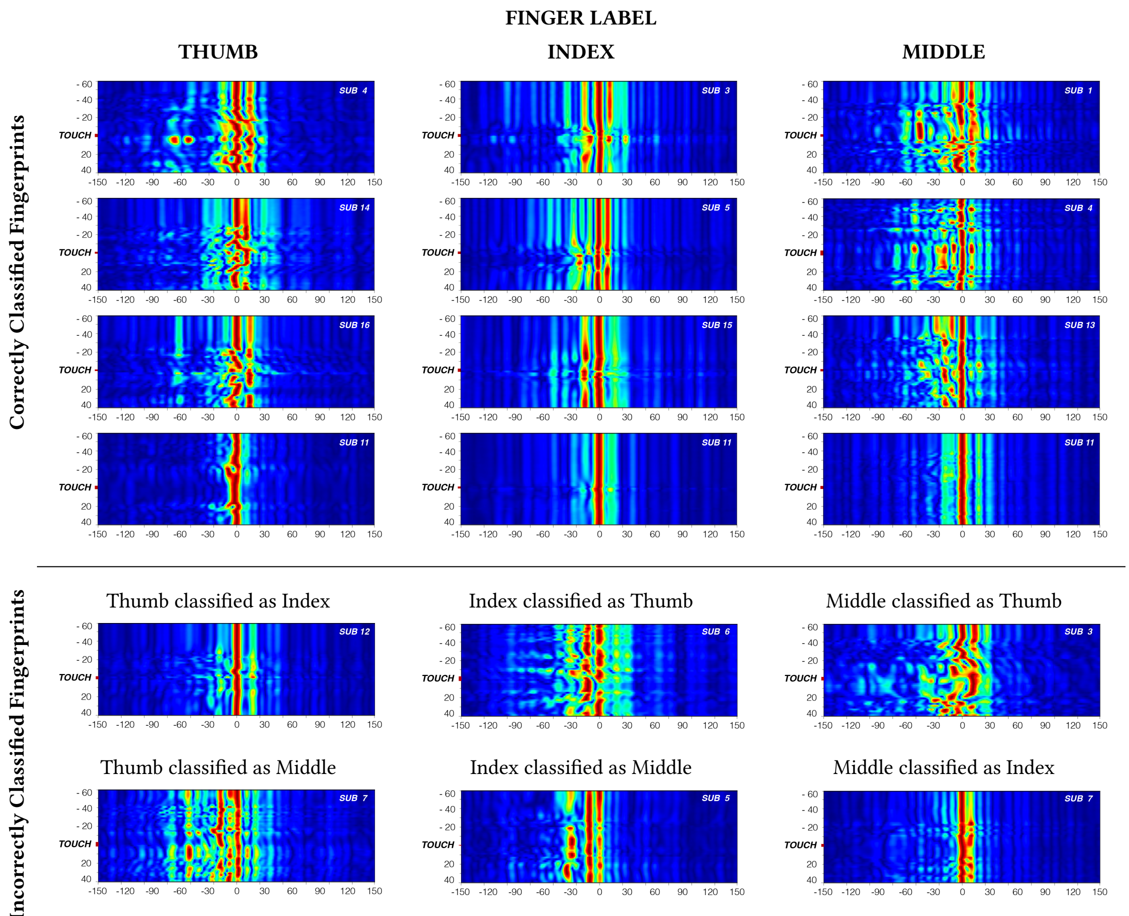

Figure 6: Representative sonar fingerprints. Columns show sonar fingerprints labeled as Thumb (left), Index (center) and Middle (right) finger. First four rows show correctly classified fingerprints, while the final two rows show incorrectly classified examples. In each figure, the x-axis depicts the number of samples (*nSamples* = 300) used, representing the maximum distance at which sonar reflections can be received, while the y-axis depicts the number of Zadoff-Chu sequences (*nSeqs* = 100) used, representing the period of time that data is captured from. Numbers indicating the participant who generated each example are marked in white text at the top right corner of each sonar fingerprint.