FruitScope: A Non-Invasive Fruit Ripeness Sensing System via Multi-Resolution FMCW Design and Acoustic Sensing

Shanmu Wang
University of California, Los Angeles
Los Angeles, USA
shanmu@cs.ucla.edu

Omid Abari
University of California, Los Angeles
Los Angeles, USA
omid@cs.ucla.edu

Abstract

Each year, large volumes of fruit are wasted due to improper storage and inaccurate quality checks, causing major economic losses. A low-cost, non-invasive assessment system that can be deployed across the supply chain would help ensure only ripe, high-quality fruit reaches consumers. Existing solutions, however, are often prohibitively expensive (>$10 k), difficult to set up, or unreliable in changing environments.

In this paper, we introduce FruitScope, the first low-cost (<$400), easy-to-deploy system for estimating fruit ripeness. FruitScope combines commodity Frequency-Modulated Continuous-Wave (FMCW) radar with acoustic sensing to achieve high accuracy and robustness. Our key innovations include (i) an advanced radar frame structure that increases commodity FMCW radar measurement granularity within limited bandwidth, (ii) a pseudo-noise method for capturing acoustic fingerprints, and (iii) a two-stage contrastive learning pipeline that refines predictions. We built an end-to-end prototype and evaluated its performance on four different fruits (avocado, mango, pear, and persimmon) and under multiple environmental disturbances. Our results show that FruitScope delivers a very low Normalized Root Mean Square Error of 2.98% to 4.63%, well below industry-accepted thresholds, while remaining resilient to environmental interferences.

CCS Concepts

• **Human-centered computing → Ubiquitous and mobile computing systems and tools**.

Keywords

Food Sensing, Multi-Modal Sensing

ACM Reference Format:
Shanmu Wang and Omid Abari. 2026. FruitScope: A Non-Invasive Fruit Ripeness Sensing System via Multi-Resolution FMCW Design and Acoustic Sensing. In *ACM/IEEE International Conference on Embedded Artificial Intelligence and Sensing Systems (SenSys '26), May 11–14, 2026, Saint Malo, France.* ACM, New York, NY, USA, 15 pages. https://doi.org/10.1145/3774906.3802794

SenSys '26, May 11–14, 2026, Saint Malo, France

ACM ISBN 979-8-4007-2309-4/26/05
https://doi.org/10.1145/3774906.3802794

1 Introduction

Fruit waste represents a major global economic problem: the United Nations estimates that nearly half of all fruit produced each year is never consumed [1]. Growers often harvest fruit before it reaches full maturity to allow time for transportation to markets. This practice creates two key challenges. First, growers must accurately assess ripeness levels to package fruits of similar maturity together. If even one overripe fruit spoils, the entire package may be discarded. Moreover, as the industry increasingly adopts robotic harvesting systems, this task becomes even more complex since robots struggle to sense ripeness with the same accuracy and subtlety as human pickers. Second, consumers need to evaluate ripeness at the point of purchase, since preferences vary depending on how soon they plan to consume the fruit. Traditional methods such as squeezing or other manual tests by growers or consumers can damage the produce [46], ultimately contributing to further waste.

Past work has proposed different techniques to enable non-invasive fruit sensing. However, each system has its limitations. For example, near-infrared (NIR) spectroscopy and sub-terahertz (sub-THz) systems achieve high accuracy [2, 3, 10, 14, 22] yet cost $10k–$150k [13, 34, 35]. Wi-Fi and millimeter-wave (mmWave) solutions are cheaper [33, 57, 58, 71] but either fragile in dynamic settings or limited by coarse measurement granularity. Vibration techniques demand costly hardware (e.g., pulsed lasers) [17, 19] or forced transducer contact [23, 36, 65], risking fruit damage and hindering deployment. To the best of our knowledge, no existing system offers the combination of low cost, practicality, accuracy, and robustness.

To fill this gap, we introduce FruitScope, a *low-cost*, *easy-to-deploy*, *multi-modal* system that estimates ripeness accurately and reliably under real-world conditions. FruitScope fuses commodity mmWave Frequency-Modulated Continuous-Wave (FMCW) radar with acoustic sensing using a small speaker and a microphone array. Our multimodal approach enables capturing fruit properties related to ripeness from different perspectives. In particular, our experiments, along with previous work, show that mmWave can penetrate the fruit flesh, and that the received-signal-strength (RSS) changes as sugar and water content evolve [58, 71, 76, 77]. Acoustic waves, meanwhile, reflect and attenuate according to mechanical properties such as softness, density, and internal texture, etc., which also change as the sugar-water content changes [63]. Although previous systems have shown the feasibility of using RF or acoustic signals for ripeness sensing, their single-modality approaches provide insufficient accuracy and adaptability. In this work, we increase the measurement diversity by combining two different perspectives, which helps increase accuracy and adapt to diverse fruit types, as different fruits can exhibit varying ripening

characteristics. However, integrating both modalities into a single system requires addressing several challenges. Below we describe each challenge and describe how FruitScope solves them with novel solutions.

C1 - ***High-granularity internal RF sensing of a fruit.*** The first challenge is to sense the fruit's interior with high-granularity using only a low-cost commodity FMCW radar. Commercial radar's bandwidth (B) is only 4 GHz which enables a finest ~4 cm measurement granularity ($c/2B$) [59]. Such a granularity is inadequate for sampling the underlying reflection profile of a small fruit (such as a persimmon), and hence cannot provide reliable ripeness estimation. To solve this problem, we propose a novel frame structure for FMCW and algorithm to dynamically adjust the FMCW frames. Our *advanced frame structure* splits a conventional radar frame into sub-frames, each transmitting chirps with different bandwidths, thereby offering measurements at shifted range-bin centers with varying granularity levels. Using our advanced frame structure and dynamically adjusting it, we achieve alternate sampling of the reflection profile of the fruit at 1 cm granularity using the same 4 GHz radar hardware. It is worth mentioning that our approach seamlessly integrates with commodity FMCW radar without requiring any hardware changes or any help from the manufacturer.

C2 - ***Force-free acoustic fingerprinting of a fruit.*** The second challenge is to acquire reliable acoustic profiles without applying a clamping force to the fruit. Prior acoustic-based systems typically require a contact pressure of 5-7 N to minimize transmission loss due to acoustic impedance mismatch at the sensor–fruit interface. The applied force not only increases the risk of bruising the fruit but also slows down the process [53]. To solve this limitation, we design a 3D-printed holder structure where the fruit can simply rest on top of a speaker, with a microphone array above it. However, due to the gap between the fruit and the speakers/microphone, some acoustic attenuation occurs. To overcome this signal attenuation and precisely track the acoustic channel changes induced by the fruit, we design to emit a pseudo-noise (PN) sequence in the 17-22 kHz ultrasonic band and correlate the recording with the known PN to estimate the acoustic channel impulse response (CIR). Such a technique functions as a matched filter and provides processing gain, which eliminates the need for forceful contact. Our experiments demonstrate that this method effectively captures acoustic fingerprints of the fruit across different ripeness stages while maintaining deployment simplicity.

C3 - ***Label-efficient data collection and learning.*** Finally, the third challenge is to reduce the manual effort required to collect high-quality ripeness labels (e.g., Brix and dry matter) for training our estimation algorithm. The golden method for collecting ripeness ground truth is destructive, labor-intensive, and expensive at scale. We therefore design a novel *label-efficient data collection pipeline* that seeks temporal and observation diversity instead of densely labeling. Specifically, we capture the natural ripening process of fruits every 12 hours over time, combined with multiple sensor views (i.e., different orientations of fruits) at each timestamp. After the evening session, only a small number of fruits are destructively tested for ripeness, greatly reducing the labor effort. Therefore, for each fruit, only a single timestamp is destructively labeled, while the rest remain unlabeled but still reflect its continuous ripeness changes. To fully leverage such a structure, we build a *two-stage learning framework*. First, we apply self-supervised contrastive pretraining that treats different views of the *same fruit at the same time* as natural positives, shaping a radar-acoustic encoder that is viewpoint-invariant yet fruit/ripeness-specific; second, we attach a lightweight regression head and fine-tune the full model on the sparse labels using a hybrid (classification + regression) loss that improves model convergence. This design attains good accuracy across fruit types while requiring far fewer labeling efforts than a conventional fully supervised approach.

We built a prototype of FruitScope using Commercial Off-the-Shelf (COTS) parts, costing less than $400 in total. With this prototype, we collected 5,400+ samples from four representative climacteric fruits, avocado, mango, pear, and persimmon (20 specimens each). To obtain the ground truth, we manually measured ripeness metrics destructively via Brix and Dry-matter tests as in the fruit industry [14, 27, 42]. With this dataset, we evaluate the estimation accuracy by conducting the *Leave-One-Out-Cross-Validation (LOOCV)*. Our results demonstrate an average Normalized Root Mean Square Error (NRMSE) ranging from 2.98% to 4.63% across these fruit types, well below the 10% food industry tolerance that ensures negligible and imperceptible impact on taste to the consumer [22]. We also evaluate the performance of the system under various environmental conditions. Our results demonstrate that our system remains stable amid dynamic environments with moving people, ambient audio noises, and co-located devices. Finally, we conduct extensive ablation studies to demonstrate the benefit of multimodal fusion and each design element.

To summarize, we make the following contributions:

- We introduce FruitScope, the first low-cost (<$400) and easy-to-deploy system for fruit ripeness estimation, while achieving high accuracy and robustness to environmental noises.
- We propose a novel advanced FMCW frame structure, boosting commercial radar's effective measurement granularity to 1 cm without requiring additional bandwidth. We also design a 3D structure and a PN-based acoustic CIR extraction method, enabling a simple force-free setup for acoustic sensing.
- We introduce a label-efficient data collection pipeline along with a two-stage learning algorithm for ripeness estimation, reducing reliance on destructive labels.
- We build a prototype of FruitScope using off-the-shelf devices, and evaluate its performance on four representative climacteric fruits.

2 Background

2.1 Fruit Categories and Ripeness Metrics

Fruit Categories. Fruits can be divided into *climacteric* and *non-climacteric* categories. Climacteric fruits, such as pears, continue ripening after harvest, unlike non-climacteric fruits such as oranges. Climacteric fruits must be harvested before fully ripening to withstand transportation, yet their ripeness continues to evolve post-harvest, making accurate assessment challenging for both consumers and the fruit industry during storage and retail. Due to this particular challenge, this work focuses on climacteric fruits.

Ripeness Metrics: Brix and Dry Matter Content. The Brix value is a metric in the food industry that measures sugar content or the sweetness of fruits. It quantifies the total soluble solids in

the fruit, with one degree Brix (1°Bx) corresponding to 1 gram of sucrose in 100 grams of aqueous solution. As fruits ripen, starches convert to sugars, causing the Brix value to increase, making it an excellent indicator of ripeness. In the fruit industry, Brix is typically measured using destructive methods that require extracting juice from a cut fruit sample and analyzing it with a digital refractometer.

Dry Matter Content (DMC) is another critical metric for determining fruit ripeness. DMC refers to all solid components of the fruit, excluding water content, including sugars, carbohydrates, oils, proteins, etc. The accumulation of these solids increases as fruits ripen. A fruit typically reaches its maximum DMC at peak ripeness. Typically, measuring DMC involves drying small fruit samples in an air-forced oven at 200°F until achieving constant dry weight, then calculating the ratio of dry matter to original weight. Because Brix and DMC are widely accepted ripeness metrics [2, 22], we focus on these as our assessment criteria.

2.2 FMCW Technology

Frequency Modulated Continuous Wave (FMCW) radar is a powerful wireless sensing technology that offers precise distance measurement and material property detection. FMCW radar transmits a signal with linearly increasing frequency $s_{TX}(t) = e^{j2\pi(f_0 t + \frac{1}{2}St^2)}$, where f_0 is the initial frequency and S is the chirp slope. When encountering a scatterer at distance R, the returning signal experiences a round-trip delay $\tau = \frac{2R}{c}$ and signal attenuation. The signal attenuation factor A depends on both distance R and the scatterer's Radar Cross Section (RCS) σ following $A \propto \frac{\sigma}{R^4}$ [44].

Mixing the transmitted and received signals produces an intermediate frequency (IF) signal with beat frequency $f_b = \frac{2SR}{c}$ that correlates with distance. By sampling this signal and applying a Fast Fourier Transform (FFT), we obtain a range profile showing received signal strength (RSS) at different distances. Since RCS is influenced by material properties, changes in the RSS profile can indicate material changes. This capability makes FMCW radar particularly valuable for diverse applications, including material sensing [50, 69, 72], liquid sensing [29, 68], motion capture [70], robotics [16], leaf wetness sensing [31, 32], and fruit sensing [38, 58].

2.3 Acoustic Channel and Pseudo-Noise Sequence

Similar to wireless signals, audio signals experience multipath propagation from a speaker to a microphone, where the received signal $r(n)$ is a superposition of multiple delayed and attenuated versions of the sent signal $s(n)$, with $r(n) = \sum_{i=1}^{N} \alpha_i \cdot s(n - \tau_i) = h(n) * s(n)$. Here, $h(n)$ is the acoustic channel impulse response (CIR). The amplitude spectrum of the channel response in the frequency domain $|H(f)|$ is related to the internal structure and physical properties of the transmission media [12]. Placing a fruit between the transceivers changes the CIR, reflecting the fruit's properties as an acoustic medium. Variations in fruit shape and ripeness produce different CIRs, enabling the extraction of structural and ripeness information.

Pseudo-Noise (PN) sequences are binary sequences that appear random but are deterministic. They have excellent auto-correlation properties, making them valuable for CIR estimation [43]. Common PN sequences include m-sequence [48], GSM training sequence [73], Kasami sequence [56], and Golay sequence [64]. More recently, PN sequences have been applied for various applications, such as gesture tracking [73], child presence detection and breadth detection [56, 75]. While previous acoustic sensing research focuses on environmental changes, our work firstly applies these techniques to sense minor changes in fruit's internal properties.

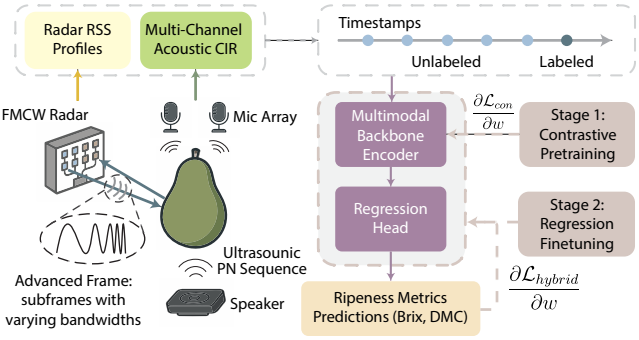


Figure 1: System overview of FruitScope. We propose novel techniques to obtain fine-grained RF properties and acoustic features of fruits. Combined with a label-efficient two-stage training pipeline, FruitScope achieves robust and accurate ripeness estimation with minimal labeled data.

3 FruitScope System Overview

FruitScope is a novel low-cost sensing system that combines FMCW radar and acoustic sensing to estimate fruit ripeness. The key idea is that RF and acoustic sensing provide different information about the fruit's internal properties, which can be fused to deliver accurate and robust estimation of the fruit's ripeness metrics such as Brix and DMC for a variety of fruits.

Figure 1 shows the system architecture of FruitScope, which consists of two main parts: (i) sensing via RF (Section 4) and acoustic modalities (Section 5), and (ii) a label-efficient data collection pipeline (Section 6) paired with the corresponding two-stage learning framework (Section 7).

4 Fine-Grained RF Sensing

Fruits are complex dielectric materials that can be approximated as a set of concentric layers (i.e., peel, flesh, and seed), each exhibiting distinct electromagnetic properties [2]. As fruit ripens, changes in water and sugar content alter the dielectric properties of each layer, leading to corresponding variations in the receive-signal-strength (RSS) profiles captured by an FMCW radar. Recent studies have demonstrated that mmWave at 60 GHz [71], 75~110 GHz [76], and even 110~170 GHz [77] radios can penetrate fruits and reveal their internal characteristics, which is a significant advantage over vision- or near-infrared-based methods.

Limitation of COTS Radars. However, commodity radars face a significant bottleneck: their limited bandwidth (typically ≤4 GHz) results in a coarse measurement granularity (every 4 cm) for the RSS profile, leading to inadequate information and increasing ambiguity. To understand this limitation deeper, consider M scatterers within the radar's field of view, each at distances R_i. Each scatterer

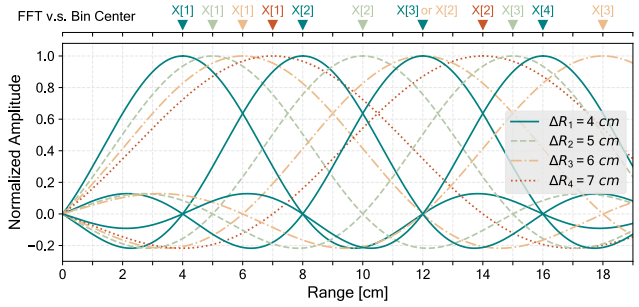


Figure 2: Examples of the amplitude of $sinc_N(\frac{R_i - k\Delta R}{N\Delta R})$ (normalized by $\frac{1}{N}$) for different sub-frames at different range bins, with $N = 256$ and measurement granularity $\Delta R = 4, 5, 6, 7$ cm, respectively.

generates an IF signal with a beat frequency f_i. The RSS measurements are obtained by sampling the IF signals at a rate of f_s over N samples and applying an N-point FFT, which yields:

$$X[k] = \frac{1}{N}\sum_{i=1}^{M} A_i e^{j\phi_i} \cdot e^{j\pi(N-1)\left(\frac{f_i}{f_s} - \frac{k}{N}\right)} \cdot sinc_N\left(\frac{f_i}{f_s} - \frac{k}{N}\right). \quad (1)$$

Here k is the range bin index, $sinc_N(x) = \frac{sin(\pi N x)}{sin(\pi x)}$ is the Dirichlet kernel induced by the finite FFT window, and A_i and ϕ_i are the attenuation and phase caused by the scatterer's reflection[1]. The frequency resolution of the FFT $\Delta f = \frac{f_s}{N}$ corresponds to a measurement granularity of $\Delta R = \frac{c\Delta f}{2S} = \frac{cf_s}{2SN} = \frac{c}{2B}$, which indicates $\frac{f_i}{f_s} = \frac{2SR_i}{cf_s} = \frac{1}{N}\frac{R_i}{\Delta R}$. Therefore, we can rewrite the RSS profile as:

$$X[k] = \frac{1}{N}\sum_{i=1}^{M} A_i e^{j\phi_i} \cdot e^{j\pi\frac{N-1}{N}\left(\frac{R_i - k\Delta R}{\Delta R}\right)} \cdot sinc_N\left(\frac{1}{N}\left(\frac{R_i - k\Delta R}{\Delta R}\right)\right), \quad (2)$$

where $k\Delta R$ is the center of the range bin associated with $X[k]$. Thus the FFT acts as a matched filter centered at each $k\Delta R$, aggregating reflections across the radar's field of view[2], with the granularity limited by ΔR. Currently most of FMCW radars offer the finest granularity of ~4 cm, resulting in coarse measurements and low accuracy as in the prior work [58].

Multi-Resolution Design. The key to addressing this limitation is to sample the reflection profile more densely. Although the finest measurement granularity is upper-bounded by the bandwidth, we can shift the sampled bin centers to achieve an effective finer granularity. To do so, we leverage the capability of FMCW radars to configure chirps differently within a frame. We use this feature to create an *advanced frame structure*, which simply repeats the matched-filter view at several ΔR_ℓ, filling the gaps and increasing the effective measurement granularity for fruit reflection patterns (Figure 2). In particular, our advanced frame consists of multiple sub-frames, with each of them configured to send out chirps of distinct bandwidth, which yields a set of separate FFT representations of the IF signals with distinct ΔR_ℓ. For instance, a single frame with $\Delta R = 4$ cm provides RSS measurements centered at 4, 8, 12 cm, etc. (corresponding to $X[1], X[2], X[3]$, etc. in Figure 2). Using multiple sub-frames with slightly different resolutions (e.g., ΔR =4, 5, 6, 7 cm) shifts the bin centers, yielding measurements centered at 4, 5, 6, 7, 8 cm, etc. (corresponding to $X[1]$, $X[1]$, $X[1]$, $X[1]$, $X[2]$, etc. in Figure 2).

Note on Range Resolution vs. Measurement Granularity. Notably, our advanced frame does not narrow the mainlobe or resolve targets closer than $\Delta R_{\text{res}} = c/2B$; instead, it increases the sampling density and offers finer information for fruit's reflection pattern, which helps improve estimations compared with single-frame FMCW methods [58], without requiring extra bandwidth or customized firmware for the hardware.

Apart from the multi-resolution design, we also exploit the time-division-multiplexing multiple-input multiple-output (TDM-MIMO) mode that most radars already support to create multiple virtual channels, which provides more detailed spatial information and boosts robustness to the fruit's surface irregularities. With all these designs, FruitScope is able to obtain fine-grained RF signatures of fruits which reveal their ripeness information.

5 Force-free Acoustic Fingerprinting

In addition to RF sensing, FruitScope incorporates acoustic signals as a second modality to provide measurement diversity. While mmWave sensing captures dielectric variations, acoustic waves are sensitive to the fruit's properties such as density and elasticity, which also evolve during ripening as the sugar water content changes [12]. Fruits are nonhomogeneous composite materials composed of organic matter, inorganic salts, and water. As they ripen, changes in internal composition lead to changes in the material's elastic modulus and density [63]. Prior acoustic-based approaches have leveraged these changes by transmitting sound waves through the fruit using acoustic device pairs, where variations in signal attenuation or phase correspond to internal changes [5, 7, 23, 24, 74]. However, such methods typically necessitate a 5-7 N force to ensure perfect transducer-fruit contact, which can potentially damage the fruit, introduce additional operation complexity, and extend the measurement time [53].

To capture reliable acoustic fingerprints of fruits without needing such contact force, we utilize a pseudo-noise (PN)-based sensing chain to estimate the acoustic channel impulse response (CIR), which provides processing gain to overcome the attenuation and gives insights into fruits' acoustic properties. In particular, we select Golay sequences [64] due to their complementary properties that reduce noise and improve CIR estimation accuracy. The Golay sequence comprises two complementary sequences, $A(n)$ and $B(n)$, with the property that their combined auto-correlation produces a sharp peak: $R_A(l) + R_B(l) = 2N\delta(l)$, where $R_A(l)$ and $R_B(l)$ are their auto-correlation results of $A(n)$ and $B(n)$ at lag l respectively, N is the sequence length, and $\delta(l)$ is the Kronecker delta function. When both sequences are transmitted through the same channel $h(n)$, the CIR can be estimated in the frequency domain as:

$$H(f) = \frac{1}{2N}\left(\mathcal{Y}_A(f)\mathcal{A}^*(f) + \mathcal{Y}_B(f)\mathcal{B}^*(f)\right), \quad (3)$$

where $\mathcal{Y}_A(f)$ and $\mathcal{Y}_B(f)$ are the FFTs of $y_A(n)$ and $y_B(n)$, while $\mathcal{A}^*(f)$ and $\mathcal{B}^*(f)$ are the conjugates of the FFTs of the original

[1]The attenuation factor A_i is characterized by the scatter's property while $\phi_i \approx 2\pi f_0 \tau$ is due to reflection path, where τ is the round-trip time between radar and the scatter

[2]Scatterers near the bin center ($k\Delta R \pm \frac{\Delta R}{2}$) dominate $X[k]$ due to the $sinc_N$ function's sidelobe decay.

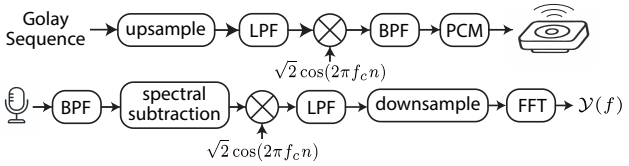


Figure 3: Modulation and demodulation scheme for acoustic sensing in the ultrasound band.

sequences. To implement this method in practice, we address two challenges.

First, to avoid human-audible disturbance, we modulate the Golay sequences to the ultrasound frequency band (17-22 kHz). We first up-sample the baseband signal to limit its bandwidth, which is done by zero-padding and low-pass filtering [40]. We then modulate it with a carrier frequency of $f_c = 19.5$ kHz. The modulated signal is bandpass filtered, PCM encoded, and played through a speaker, traveling through the fruit to reach the microphones (Figure 3 up).

The other challenge in using acoustic for fruit sensing is that the received acoustic signals contain significant noise from hardware imperfections and environmental sources. To address this, we first apply bandpass filtering to remove out-of-band noise upon receiving. However, we observed a stationary narrow-band noise at 10.8 kHz whose harmonic at 21.6 kHz falls within our desired frequency band, caused by clock coupling. To eliminate this in-band stationary noise (with relatively stable frequency content and amplitude), we apply spectral subtraction [6, 67] using short-time Fourier transform (STFT) with a window size of 512 samples. We estimate the noise profile by averaging the power spectrum of the first 20 frames and then subtract it from subsequent frames. After denoising and demodulation, we obtain the frequency-domain CIR estimation $H(f)$ by processing both Golay sequences according to Equation 3 (Figure 3 down). The magnitude $|H(f)|$ serves as the acoustic profile of the fruit.

Such a design allows our system to estimate repeatable acoustic fingerprints that successfully track the fruit's ripeness changes, while maintaining a minimal setup requirement. Since during ripening, different fruits can exhibit varying characteristics (some may show more significant changes in RF signatures, while others in acoustics), our multimodal approach could enhance the system's applicability across a broader range of fruit types.

6 Data-collection Pipeline

Now that the sensing principles are established, we turn to building the dataset that trains our model. The main challenge for this is that golden standard ripeness tests (Brix and dry matter content) for fruits are destructive, labor-intensive, and costly at scale. Therefore, instead of seeking dense annotations, we propose a *label-efficient* strategy that captures rich measurements over time and across viewpoints, while collecting ground-truth labels only sparsely.

To enable this, we design a 3D-printed, COTS-based data collection platform, as shown in Figure 4. With a mmWave FMCW radar positioned beside the specimen, a miniature speaker mounted underneath, and a microphone array above, the holder on our platform allows us to rotate or tilt the fruit, yielding multiple *views*, i.e.,

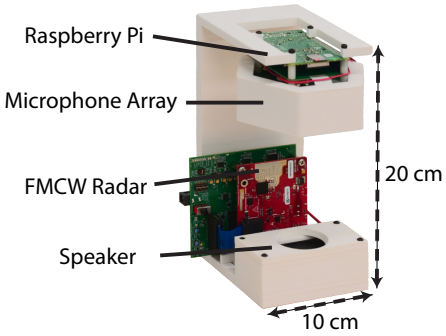


Figure 4: Data collection platform.

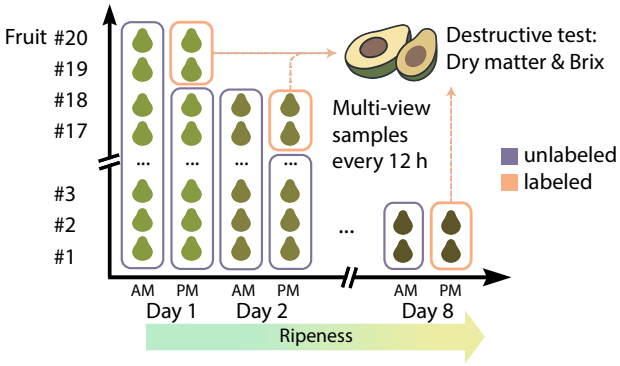


Figure 5: Data collection pipeline of FruitScope. We collect data from fruits at different views and timestamps, which provides *aspect diversity* and *temporal diversity* to the dataset, while keeping manual labeling to a minimum.

signals from the same fruit at different orientations with respect to both sensors.

With the hardware in place, data collection proceeds in a simple and efficient way as illustrated in Figure 5. First, we purchase fruits early in the ripening cycle (e.g., light-green avocados, firm mangoes, etc.) so that each specimen continues to ripen. Every 12 hours we record *ten* views per fruit. For clarity, we define "different views of the fruit" as signals captured from the same fruit at the same timestamp but with different orientations or tilt angles relative to the radar and acoustic sensors. After the evening session, we destructively measure just *two or three* fruits for ripeness metrics. After eight days, each specimen has been labeled at a certain timestamp; subsequent measurements are impossible for those tested fruits, but all earlier (unlabeled) readings still capture their ripeness progression.

At the end, for each fruit specimen we obtain (i) a *multi-view, labeled* set at one timestamp, and (ii) *multi-view, unlabeled* data at all earlier timestamps. This design dramatically reduces labeling effort yet supplies rich variations in both aspect (views) and temporal (ripening) dimensions to train a robust, generalizable model.

7 Ripeness Prediction

Next, we present our design for extracting and fusing features from the two sensing modalities and a two-stage learning framework for fully leveraging our dataset structure.

7.1 Backbone Encoder and Fusion

With our sensing design, both sensing streams are converted to matrices $\mathbf{X} \in \mathbb{R}^{C\times N}$: where C represents TX–RX pairs (radar) or microphones (acoustic), and N represents range bins or frequency bins. This exposes spatial and spectral patterns along two axes.

We extract features using separate but architecturally identical CNN branches, each with two sequential 1-D convolutions that decouple spectral and spatial features:

- **Spectral block.** A Conv1D(k, kernel = $(1, m)$, stride = $(1, s)$) learns RSS/frequency patterns per channel.
- **Spatial block.** A Conv1D(k, kernel = $(C, 1)$) then aggregates across the C channels, capturing inter-antenna or inter-microphone correlations.

Each convolution is followed by BatchNorm, activation, and dropout. Features are flattened and projected to a d-dimensional embedding for further processing: $\mathbf{h}_r, \mathbf{h}_a \in \mathbb{R}^d$. Such a scheme has shown its effectiveness in multi-channel signals with similar characteristics [54, 55].

Since different fruits respond unequally to radar and acoustic probes, we let the network adaptively fuse features. To do so, we compute a summary vector $\mathbf{u} = \mathbf{h}_r + \mathbf{h}_a$ and apply a two-layer MLP with *tanh* activations. We apply softmax to the output to obtain weights for each modality, and finally fuse the two modalities with a weighted-sum mechanism:

$$\mathbf{s} = \mathrm{MLP}(\mathbf{u}) \in \mathbb{R}^2, \quad \boldsymbol{\alpha} = \mathrm{softmax}(\mathbf{s}), \quad \mathbf{h}_{\text{fused}} = \alpha_r\, \mathbf{h}_r + \alpha_a\, \mathbf{h}_a,$$

where $\alpha_r + \alpha_a = 1$. This weighted-sum mechanism is a common practice for adaptive fusion [8, 18, 41], yielding a single fused representation $\mathbf{h}_{\text{fused}}$ for further regression.

7.2 Two-stage, Label-efficient Learning

Building upon the backbone encoder and fusion mechanism, we propose a two-stage learning framework to leverage our data structure effectively. As mentioned, our dataset is highly *imbalanced*: each fruit contributes a single, multi-view *labeled* timestamp, and many earlier *unlabeled* timestamps that record its ripening trajectory. To exploit this structure, we propose a two-stage strategy: (i) contrastive pre-training on the unlabeled portion and (ii) supervised fine-tuning with a hybrid regression loss on the sparsely labeled subset.

Stage 1: contrastive pre-training on unlabeled data. Let the unlabeled set be $\mathcal{D}_{\text{unlabeled}} = \left\{ x_i = (\mathbf{r}_i, \mathbf{a}_i) \right\}$, where each sample x_i contains an RF feature $\mathbf{r}_i$, an acoustic feature $\mathbf{a}_i$, and two tags: fruit ID f_i and timestamp t_i. Unlike conventional contrastive learning methods that rely on artificial augmentations [9], our data offer *natural* positives: Specifically, within a mini-batch of size N, we define

$$\mathcal{P}(i) = \left\{ j \neq i \mid f_i = f_j, \text{and } t_i = t_j \right\}, \quad \mathcal{N}(i) = \{ j \neq i \} \setminus \mathcal{P}(i),$$

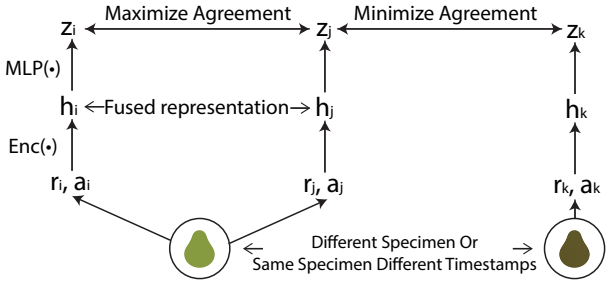


Figure 6: The contrastive pretraining framework.

i.e., *positives* are different views from the *same fruit measured at the same timestamp* and therefore corresponds to the same ripeness level, while all remaining samples serve as *negatives*.

As shown in Figure 6, each x_i passes through the encoder and a projection head to yield a latent $\mathbf{z}_i$. We compute the NT-Xent loss as follows [9]:

$$\mathcal{L}_{\text{NTX}} = -\frac{1}{N} \sum_{i=1}^{N} \log \left(\frac{\sum_{j\in\mathcal{P}(i)} \exp\left(\mathrm{sim}(i,j)/\tau\right)}{\sum_{k\in\mathcal{P}(i)\cup\mathcal{N}(i)} \exp\left(\mathrm{sim}(i,k)/\tau\right)} \right), \quad (4)$$

where $\mathrm{sim}(i,j) = \frac{\mathbf{z}_i^T \mathbf{z}_j}{\|\mathbf{z}_i\| \|\mathbf{z}_j\|}$ is the cosine similarity and τ is the temperature parameter. This objective forces the representation to be invariant to viewpoint yet sensitive to fruit identity and ripeness state.

Stage 2: supervised fine-tuning with hybrid regression loss. After pre-training, we attach a lightweight regression head and fine-tune the entire network on the labeled subset $\mathcal{D}_{\text{labeled}} = \left\{ (x_i, y_i) \right\}$, where y contains the ripeness metrics. To balance *precision* and *flexibility*, we adopt a hybrid objective that blends classification over coarse bins with continuous regression [25]:

$$\mathcal{L}_{\text{hybrid}} = \lambda\, \mathcal{L}_1 + (1-\lambda)\, \mathcal{L}_{\text{cls}}. \quad (5)$$

Specifically, we discretize the target range into C bins, predict a softmax over bins, and calculate a cross-entropy loss ($\mathcal{L}_{\text{cls}}$). We then use the softmax probabilities to compute the expectation of bin centers as the final predictions, and refine that value with an ℓ_1 loss ($\mathcal{L}_1$). Empirically, we found that $\lambda = 0.5$ yields the best trade-off between flexibility and precision and improves model convergence.

8 Implementation

This section details the implementation of FruitScope, including hardware components and system configurations.

Hardware design: Figure 4 shows our prototype built with COTS devices. The core hardware includes a Texas Instruments AWR1843 mmWave radar (77-81 GHz) with 3 transmitters and 4 receivers creating 12 virtual channels, and a ReSpeaker 6-microphone array [49], a 3W speaker, and a Raspberry Pi for audio processing. We designed a 3D-printed structure positioning the radar 5 cm from the fruit, the speaker directly beneath, and the microphone array 15 cm above. This arrangement ensures acoustic signals travel through the fruit before capture.

Cost Analysis: The total hardware cost of our current prototype is $395. This includes $300 for the radar module, $40 for the microphone array, $50 for the Raspberry Pi, and $5 for the

speaker. However, this cost can be significantly reduced through mass production. Note that although the current prototype includes a DCA1000EVM [62] to capture raw IF data, it is for understanding the signal properties better and designing our algorithms. FruitScope do not require it to estimate the fruit ripeness as the radar board itself can output the range-profile, which is sufficient for our estimation algorithm.

System Configurations: For RF modality, we configure four sub-frames with measurement granularity of 4, 5, 6, 7 cm, respectively. We use the mmWave sensing estimator tool [61] provided by TI to get the chirp configurations. Each sub-frame repeats chirps 12 times per TX, with results averaged during post-processing. For each chirp, we sample 256 points and collect RSS data from range bins between 4–16 cm. For acoustic modality, we operate at a 48 kHz sampling rate, generating 512-sample Golay sequences modulated to 17-22 kHz. Sequences repeat 3 times for measurement consistency. The Raspberry Pi captures audio as numpy arrays and transfers files to PC via SFTP. We use Butterworth bandpass filtering and spectral subtraction before using the acoustic signals for CIR estimation.

Model Architecture and Training: We feed two sensing matrices into a shared 64-dimensional feature space. The radar encoder uses (1,4) kernels with stride (1,1) and ReLU activation, while the acoustic encoder employs (1,6) kernels with stride (1,2) and ELU activation [11]. Both use 0.2 dropout probability. The regression head consists of a two-layer network with 16 hidden units and ReLU activation. We use Adam optimizer with learning rates of 0.015 (pretraining) and 0.001 (fine-tuning), cosine annealing scheduling, and train for 1,000 pretraining epochs and 100 fine-tuning epochs with 8:2 train-validation split and early stopping.

9 Evaluation

To evaluate FruitScope, we selected four distinct fruits (avocado, mango, pear, and persimmon) based on their diverse geometry, peel structures, and internal seed characteristics, as well as their prevalence in prior work [2, 22]. We collected an eight-day dataset comprising over 5,400 multimodal measurements from 80 fruits (20 per species), purchased from different stores on different days. Of these, 800 samples include destructive ground-truth values for dry-matter content (DMC) and Brix[3]; the rest are unlabeled temporal measurements used for contrastive pre-training. We used a lightweight vision pipeline to segment each fruit from photos to show their morphological/geometry diversity[4]. We computed two descriptors: *roundness* (circularity) and *area* (pixel count). Figure 7 shows that our dataset spans a broad range on both metrics. All specimens are stored and measured at room temperature (21°C) in a typical laboratory. While our experiments were conducted in a typical indoor setting, we did our best to encompass real-world variability. For example, our dataset includes environmental noise (e.g., people were around, talking and moving, and coexisting wireless signals) and placement variance (i.e., 10 different fruit orientations per measurement time). Table 1 summarizes our DMC and Brix statistics. Although the numeric ranges appear narrow, even a two-degree increase in Brix (e.g., 15⟶17) produces a noticeably sweeter taste in mangoes.

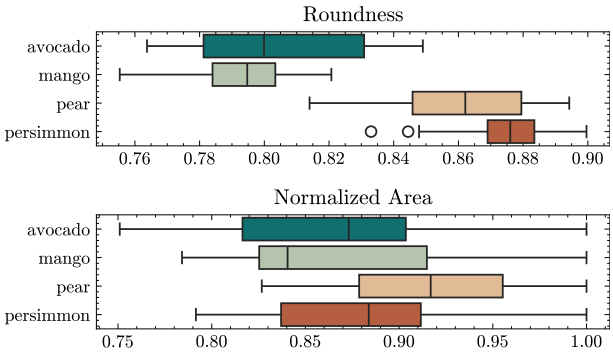


Figure 7: The boxplot of the morphology metrics showing the quartiles of the distribution.

Table 1: Statistic of the ripeness metrics.

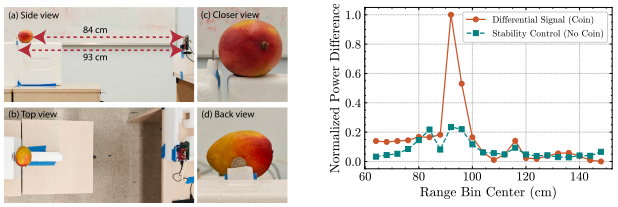
Fruit	Metrics	Mean±Std	Min	Max
Avocado	DMC (%)	25.50±1.26	23.2	27.8
Mango	DMC (%)	19.11±1.31	16.6	21.4
	Brix (°)	17.09±1.16	14.8	18.7
Pear	DMC (%)	14.67±1.00	13.3	16.8
	Brix (°)	13.35±0.81	12.1	14.9
Persimmon	DMC (%)	18.66±0.81	17.4	20.1
	Brix (°)	16.73±0.69	15.6	17.7

Our evaluation begins with feasibility experiments to validate the penetration ability of COTS mmWave radars and observe how radar and acoustic profiles change as fruit ripens. We then report the overall accuracy of FruitScope, compare single- and multi-modality configurations, and analyze robustness to environmental noise and impact of different hardware settings.

9.1 Feasibility Experiment

9.1.1 MmWave Penetration Validation. Although prior work has shown that mmWave signals at 75–170 GHz can penetrate fruit flesh and reveal internal structure [76, 77], we additionally validate this capability on COTS 77–81 GHz FMCW radars. We employ the AWR1843 and the setup in Figure 8(a). The radar and fruit face each other on dry Styrofoam risers, which are effectively transparent at mmWave and introduce negligible reflections/attenuation [39, 60]. We acquire three recordings in immediate succession without touching or moving the fruit: (i) *fruit only*, (ii) *fruit+coin* in which a U.S. quarter (compact, high-RCS reflector) is placed directly behind the fruit with a standoff < 1 cm, and (iii) *fruit only* again. The fruit and coin rest on separate cabinets so that inserting/removing the coin does not disturb the fruit. This third recording is to verify that the fruit and fixtures (risers, mounts) remain static throughout. The fruit–radar distance is 84 cm; the coin lies at 93 cm. The radar transmits 3.75 GHz bandwidth chirps, yielding a resolution of 4 cm.

[3]For avocado, only DMC is measured due to low sugar content.
[4]Side-view for avocado and mango; top-down for pear and persimmon.



(a) Experiment setup with a hidden coin reflector behind a fruit (b) Normalized power difference of the differential measurements.

Figure 8: mmWave penetration experiment.

We average beat signals across chirps and compute the complex range FFT for each recording. We then form a differential profile $S_{\text{coin}} - S^{(1)}_{\text{fruit}}$ to suppress static returns (antenna leakage, fixtures, front surface) that are common to both shots. Therefore, the magnitude of the residual highlights new energy attributable to the hidden reflector. As a reference, we also compute $S^{(2)}_{\text{fruit}} - S^{(1)}_{\text{fruit}}$.

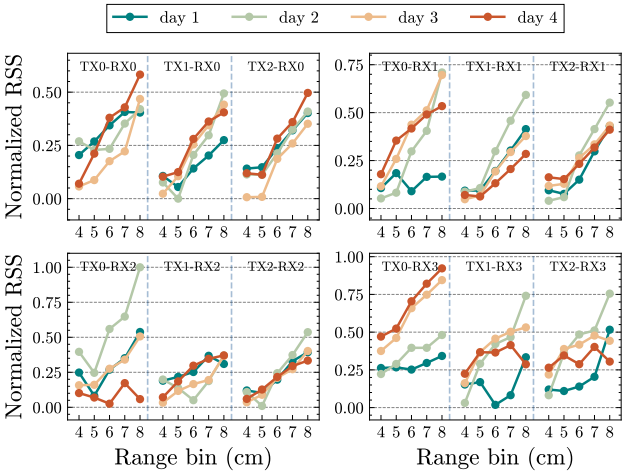
Figure 8(b) shows that the *fruit+coin* subtraction (red) exhibits a clear peak at 92–96 cm, whereas the *fruit vs. fruit* subtraction (green) shows no such feature. Because the direct line of sight (LoS) from the radar to the coin is obstructed by the fruit body, this residual peak indicates a penetration path, thereby demonstrating through-fruit propagation with the COTS automotive radar.

9.1.2 Change of Signatures during Ripening. Next, we verify that fruit ripening alters both the radar received signal strength (RSS) profiles and the acoustic channel impulse response (CIR) profiles. To do so, we monitored a single fruit for 100 hours. Over the measurement, the fruit remained stationary, and the ambient temperature was held constant. Figure 9 shows the results from a mango and yields four key observations:

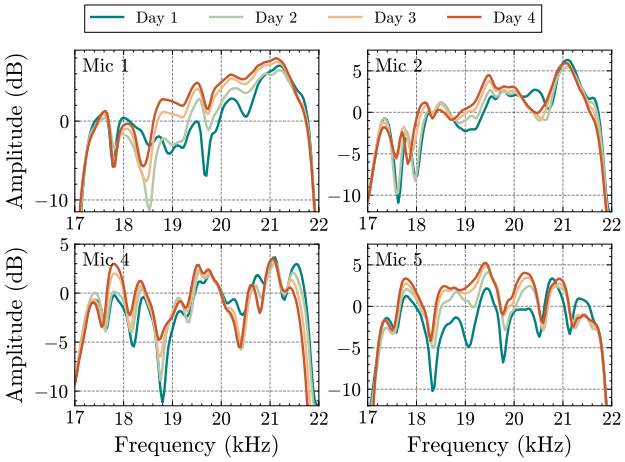
(1) **Spatial diversity in RF.** RSS traces from the 12 TX–RX paths differ markedly. Some paths rise monotonically (e.g., TX0–RX3), some rise and then fall (e.g., TX1–RX1), and a few remain nearly unchanged (e.g., TX2–RX2). Because each path probes a slightly different trajectory through the fruit, multi-channel capture reveals heterogeneity that any single path would miss.
(2) **Complex RSS evolution.** RSS values at different range bins change at different rates over time, hinting at layer-specific processes such as water loss and sugar redistribution. Finer range resolution is therefore crucial.
(3) **Acoustic sensitivity to ripening:** Generally, the magnitude of the CIR $|H(f)|$ rises gradually across the 17-22 kHz band, suggesting lower attenuation as the fruit loses water and starch converts to sugar.
(4) **Microphone-level variation:** Different microphones record distinct $|H(f)|$ curves, and their evolution spans the whole bandwidth. A multi-mic array plus a wideband PN is therefore essential for capturing full acoustic profiles.

9.2 Overall Performance

With collected data, we evaluate the overall performance of FruitScope in estimating ripeness metrics. To measure the generalizability of our method, we conduct *Leave-One-Out Cross-Validation* (LOOCV) within each species. For each fold, one fruit is held out entirely for testing. We then pre-train the model using all unlabeled data from



(a) Normalized RSS at different bin centers from different TX-RX path



(b) Amplitude of $|H(f)|$ from different microphones at different days

Figure 9: The change of the normalized RSS and acoustic channel profile $|H(f)|$ of a mango over time.

the remaining 19 fruits. Subsequently, we fine-tune the model on the labeled data from those same 19 fruits before evaluating it on the held-out test fruit. This ensures the test fruit remains *completely unseen* during all training stages. This process is repeated for each fruit specimen in the dataset and the metrics are averaged.

To assess the prediction accuracy, we report common regression metrics, including Mean Absolute Error (MAE), Standard Deviation (STD), Pearson Correlation Coefficient (Corr.), Root Mean Squared Error (RMSE), and Normalized RMSE (NRMSE). The NRMSE, expressed as a percentage of the mean measured values, is considered acceptable in the fruit industry if it is below 10% as such errors do not impact taste and are imperceptible to consumers [22]. For NRMSE, we also computed the 95% Confidence Interval (CI) of the average value and report it as margins of error (Mean±CI) to provide a measure of the statistical significance. To demonstrate the value of our label-efficient data collection protocol and learning

Table 2: Overall performance of FruitScope.

Metrics	Dry Matter Content					Brix Value				
	MAE↓	STD↓	Corr.↑	RMSE↓	NRMSE↓ [95% CI]	MAE↓	STD↓	Corr.↑	RMSE↓	NRMSE↓ [95% CI]
One Stage: Full Supervised with Labeled Dataset										
Pear	0.70	0.44	0.74	0.78	5.23% [±1.19%]	0.60	0.38	0.68	0.67	4.97% [±1.01%]
Mango	0.87	0.60	0.73	0.99	5.21% [±1.20%]	0.75	0.54	0.77	0.86	5.11% [±1.23%]
Persimmon	0.54	0.34	0.70	0.62	3.32% [±0.81%]	0.44	0.29	0.72	0.51	3.06% [±0.80%]
Avocado	0.74	0.46	0.81	0.83	3.26% [±0.87%]			\		
Two Stage: Contrastive Pretraining + Finetuning										
Pear	0.61	0.40	0.84	0.70	4.73% [±0.99%]	0.53	0.34	0.81	0.60	4.53% [±0.81%]
Mango	0.74	0.41	0.84	0.83	4.35% [±1.22%]	0.64	0.37	0.87	0.72	4.29% [±1.06%]
Persimmon	0.49	0.33	0.76	0.56	3.05% [±0.83%]	0.43	0.27	0.71	0.49	2.97% [±0.82%]
Avocado	0.66	0.44	0.89	0.76	2.98% [±0.84%]			\		

↑ means the higher the better, while ↓ means the lower the better.

framework, we also evaluate a baseline one-stage method, where the model is trained from scratch using only the labeled data.

As Table 2 shows, both methods achieve satisfactory performance (NRMSE below 10%) for all fruit types and ripeness metrics, with the two-stage method outperforming the one-stage method in all cases. Specifically, the two-stage approach consistently lowers NRMSE by 2.9-16.5% and the estimation variance (STD) by up to 31.7%, confirming that exploiting unlabeled data is key to FruitScope's accuracy and stability. For example, using the two-stage approach, the NRMSE values are 2.98%, 3.05%, 4.35%, and 4.73% for the DMC of avocado, persimmon, mango, and pear, respectively. Additionally, the low standard deviation of predictions demonstrates the consistency of our method; the two-stage approach further improved this consistency.

In comparison to existing solutions, spectroscopy [10, 14, 47] and sub-THz systems [2, 22] are often cost-prohibitive for consumer use. Meanwhile, acoustic-only methods can be too bulky for practical deployment [17, 19, 66] or lack sufficient accuracy [23], and many RF-only systems remain sensitive to environmental noise [33, 57] (See more comparison on Section 9.4), while FruitScope achieves a balance among cost-effectiveness, ease of deployment, and sensing accuracy.

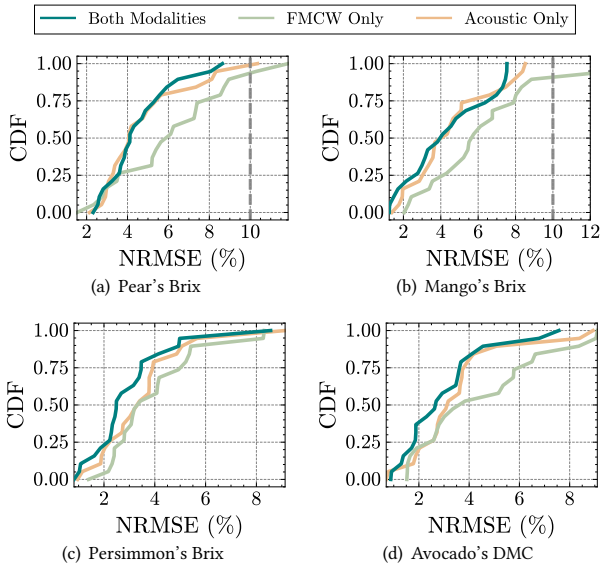


Figure 10: FruitScope's performance using single and multi-modalities. We compare the performance of the full system with using only FMCW and using only Acoustic. The predictions from multi-modality are generally more accurate than single-modality.

9.3 Benefit of Multimodal Sensing

Because radar and acoustics probe fruit properties from different perspectives, FruitScope should outperform either modality alone. We verify this by comparing three configurations: RF-only, Acoustic-only, and RF+Acoustic. In single-modality baselines, the fusion block is replaced with a two-layer MLP.

The results in Figure 10 clearly demonstrate that combining both modalities consistently produces lower NRMSE values compared to either individual modality, confirming that the integration of RF and acoustic data enhances ripeness estimation accuracy. For example, for pears both RF-only and acoustic-only show NRMSE exceeding 10% in some cases, whereas combining two modalities reduced these errors. Similarly, for mangoes and avocados, combining acoustic modality with RF modality helps improve the accuracy in general and reduces the largest errors. Lastly, for persimmons, both modality shows similar performance while their fusion yields better results.

To further investigate these relationships, we examine the weights assigned to each modality during the fusion process, which always sum to one. The visualization in Figure 11 shows different patterns of reliance on different modalities: while the acoustic modality generally receives higher attention weights across all fruit types in our dataset, avocados, mangoes, and pears show higher attention weights for the acoustic modality compared to persimmons, which aligns with our observations before. Despite the prominence of acoustic sensing in the attention weights, it is worth noting that the RF modality remains crucial for comprehensive ripeness estimation. When integrated with acoustic sensing, the RF modality helps

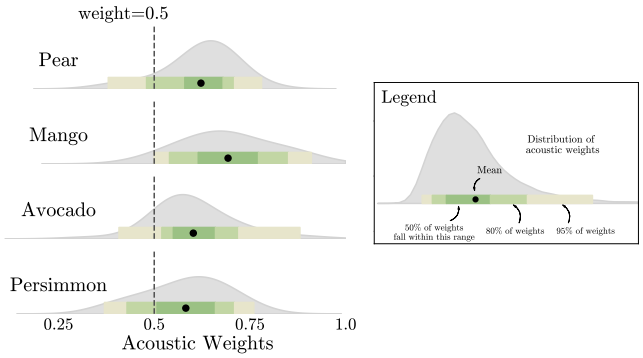


Figure 11: The Kernel Density Estimate (KDE) plot of the attention weight of acoustic modality for different fruits.

reduce ambiguity in certain cases, leading to lower overall NRMSE compared to using acoustic sensing alone. This effect is particularly notable for pears, mangoes, and avocados. These findings demonstrate that the combination of sensing technologies provides more robust and accurate ripeness assessment and helps adapt to diverse fruit types.

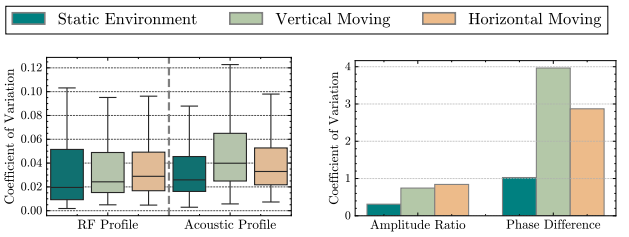
9.4 Robustness to Environment Noises

A practical advantage of FruitScope is its robustness to real-world interference, owing to the inherent resilience of both FMCW and Pseudo-Noise signals. We evaluate the system's performance under various environmental conditions, including dynamic environments, ambient noise, and mutual interference from another device.

9.4.1 Dynamic Environment. To evaluate robustness under dynamic conditions, we compare FruitScope against Wi-Fruit [33], a low-cost RF-based baseline that shares the closest operating principles to our system. Wi-Fruit predicts fruit ripeness by analyzing Wi-Fi Channel State Information (CSI) changes when fruit is placed between transceivers. Despite employing several robustness techniques such as calibration using fruit-absent measurements, phase denoising via antenna differencing, and selective subcarrier analysis, Wi-Fi CSI remains inherently sensitive to environmental dynamics.

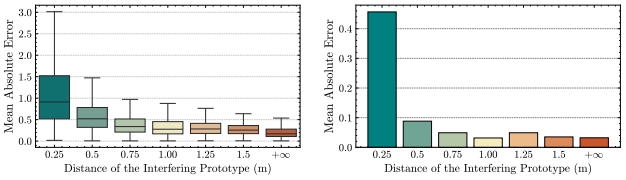
We assess dynamic environment robustness by measuring a stationary pear under two conditions: (1) a controlled *static* environment, and (2) a *dynamic* environment where an experimenter paces laterally and vertically within a 0.5 m radius. Each condition involves 10 trials. Signal stability is quantified using *coefficient of variation*: CV = STD/Mean. For FruitScope, CV is computed over trials for each range/frequency bin of the radar and acoustic profiles. For Wi-Fruit, we collect 10 measurements before and after fruit placement using an ASUS RT-AC86U router with CSI extraction via Nexmon [15] and a laptop for traffic generation. Following the original Wi-Fruit processing pipeline, we extract amplitude ratio and phase difference features, yielding 100 trial combinations over which CV is calculated.

Figure 12 demonstrates that FruitScope exhibits negligible CV growth under dynamic conditions compared to static environments. In contrast, Wi-Fruit shows fundamental limitations: even in static



(a) The Coefficient of Variation of FruitScope's raw signal (b) The Coefficient of Variation of Wi-Fruit's processed features.

Figure 12: Evaluation results under static and dynamic environments.



(a) The MAE value of the raw acoustic profile (b) The MAE value of the predictions

Figure 13: Evaluation on the mutual interference from another prototype. +∞ means no interference.

conditions, its minimal CV for amplitude ratio and phase difference is 3-8× higher than FruitScope's values. Under dynamic conditions, the CV is even 8-33× higher, indicating severe signal instability and making the ripeness estimation unreliable.

To quantify the impact of prediction accuracy, we treat static environment predictions as reference values and compute Mean Absolute Error (MAE) for dynamic conditions. For our system, the resulting MAE remains <0.10 for both DMC and Brix measurements, confirming negligible impact on estimation accuracy. These results conclusively demonstrate that FruitScope maintains both signal stability and prediction accuracy in the presence of human movement, a critical advantage for practical deployment scenarios.

9.4.2 Ambient Noise. Next, we evaluate ambient noise impact by recording sounds from various locations using a smartphone and overlaying them onto acoustic recordings taken in quiet conditions. Test environments included: (1) fruit sections of three different supermarkets (noise levels 45 dB(A), 50 dB(A), and 55 dB(A)), (2) a kitchen with both exhaust fan and air conditioner operating (60 dB(A)), and (3) an industrial warehouse (75 dB(A)). For each noise level, we create 20 variations by randomly sampling ambient noise periods and overlaying them onto the original signals. The processed acoustic profiles remain nearly identical to the originals across all conditions. Using noise-free recordings as reference, mean absolute errors after noise injection are minimal: 0.0018 (supermarket), 0.040 (kitchen), and 0.0089 (factory). This exceptional noise resilience stems from two design elements: (1) our sensing operates primarily in the ultrasound band while ambient noise exists below 10 kHz, and (2) PN sequences provide superior signal integrity through low cross-correlation with ambient sounds and high auto-correlation properties, effectively preventing external interference.

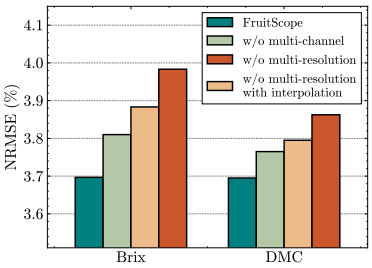


Figure 14: The impact of different RF configurations.

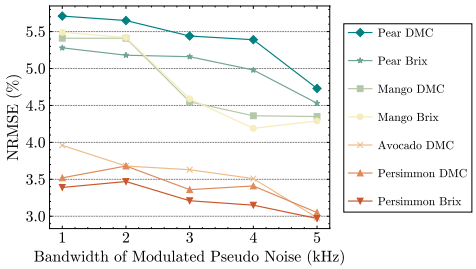


Figure 15: The impact of different bandwidths of PN sequence.

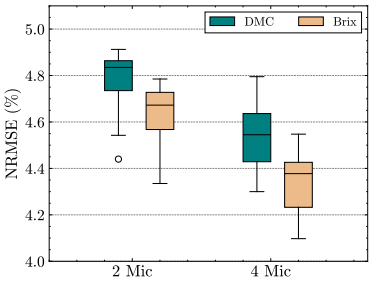


Figure 16: The impact of the number of microphones.

9.4.3 Multi-Device Compatibility. Finally, we evaluate the potential for mutual interference when multiple devices operate in close proximity, a practical concern for deployment scenarios like supermarkets where several units might assist customers simultaneously. Due to the directional nature and high attenuation of mmWave signals, RF interference is unlikely unless transmitters directly face receivers of other devices. Additionally, our radar's transmission pattern is brief and intermittent, with each measurement consisting of 144 chirps (4 subframes×3 transmitters×12 repetitions) and each chirp lasting ~250 μs, resulting in only 36 ms total transmission time per measurement. This creates an extremely low probability of temporal overlap between devices.

Since our acoustic system uses omnidirectional microphones and transmits signals for longer periods (~0.5 sec per measurement), we focus interference evaluation on the acoustic component. Our PN sequence measures approximately 30 dB(A) without fruit present, which is a relatively low level that attenuates quickly with distance. To quantify potential interference effects, we place a speaker of the same model continuously playing the PN sequence at varying distances from our device while measuring a stationary pear. For each distance, we collect 10 acoustic profiles and compare them to interference-free profiles using MAE. Results in Figure 13(a) show MAE values remain stable when the interfering source is placed 1 m or more away, indicating minimal impact on raw acoustic profiles with sufficient separation. Assessment of prediction accuracy (Figure 13(b)) also shows MAE values below 0.05 for distances ≥1 m. These findings confirm that multiple devices can operate effectively in the same environment with a minimum 1-meter separation[5], enabling scalable deployment across grocery stores, supermarkets, or farm warehouses without performance degradation from device interference.

9.5 Impact of System Configurations

We evaluate the importance of different system configurations, including RF design choices, acoustic bandwidth, and microphone array size.

9.5.1 RF Configuration. To assess the impact of our RF design, we compare our full system against three variants: (1) without multi-channel: single-channel estimation using summed power from all TX-RX pairs, (2) without multi-resolution: simple frame design with 4 cm granularity, and (3) without multi-resolution but with linear interpolation to achieve 1 cm granularity. Figure 14 shows that both multi-channel estimation and advanced frame design significantly enhance performance. Multi-channel estimation reduces average NRMSE by 1.86% (DMC) and 2.97% (Brix). More importantly, our advanced frame design decreases NRMSE by 7.20% (DMC) and 4.34% (Brix). While linear interpolation improves performance over the basic configuration, it remains less effective than our advanced frame design. This difference stems from the multi-layered, non-uniform internal composition of fruits, which linear interpolation cannot accurately represent. Our advanced frame design effectively captures reflected power at various range bins, providing a more accurate profile of the fruit's internal structure.

9.5.2 Bandwidth of Pseudo-Noise Sequence. We simulate different PN sequence bandwidths by adjusting the receiver's bandpass filter. Figure 15 shows that increasing bandwidth from 1 kHz to 5 kHz generally reduces NRMSE values. For example, mango DMC estimation NRMSE decreased from 5.41% (1 kHz) to 4.35% (5 kHz). This trend suggests that a wider bandwidth improves the system's accuracy. Therefore, we believe that utilizing higher-quality speaker and microphones with greater bandwidth could further enhance system performance.

9.5.3 Number of Microphones. We evaluate different microphone configurations by limiting data to combinations of 2 and 4 microphones from our 6-microphone array. Results averaged across all fruit types show that increasing microphone count significantly enhances performance (Figure 16). This improvement stems from increased acoustic data diversity, allowing better feature capture for ripeness estimation. Moreover, we find that optimal microphone combinations vary by fruit type. These differences likely relate to fruit shapes, which highlights the importance of using multiple microphones to capture comprehensive acoustic information and increase adaptability across fruit types.

9.6 Measurement Time

Finally, we evaluate FruitScope's end-to-end measurement latency, which is critical for practical deployment in commercial environments. Since RF and acoustic sensing operate independently, data collection and processing execute in parallel. Over 100 trials, RF data collection requires 1.98±0.02 seconds and acoustic collection takes 3.00±0.06 seconds. Signal processing completes in 0.26±0.05 and 0.15±0.02 seconds respectively, with machine learning inference requiring <0.1 seconds with an Intel i7-1165G7 CPU. In total, our

[5]One might use orthogonal PN sequences to further lower these spacing requirements.

Table 3: Summary of the related work.

Specifications	Wi-Fi Based		mmWave Based		Sub-THz Based		Vision Based	Spectroscopy	Vibration			mmWave+Sound
	FruitSensing [57]	Wi-Fruit [33]	mmWave [71]	FMCW [58]	AgriTera [2]	Meta-Stickers [22]	Cameras [4, 26, 37, 45]	NIR [10, 14, 47]	Free vibration [17, 19, 66]	Sound [36, 65]	Sound [23]	Ours
Small form factor	✓	✓	✗	✓	✓	✓	✓	✓	✗	✗	✓	✓
No need for stickers on fruits	✓	✓	✓	✓	✗	✓	✓	✓	✓	✓	✓	✓
Robust to dynamic environment	✗	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Non-intrusive to humans	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✗	✓
Brix value estimation	✗	✓	✓	✓	✓	✓	✗	✓	✓	✓	✗	✓
Dry-matter content estimation	✗	✓	✗	✗	✓	✓	✗	✓	✓	✓	✗	✓
Internal properties	✓	✓	✓	✓	✓	✓	✗	✓	✓	✓	✓	✓
Accuracy[a]	Low	High	Medium	Medium	High	High	Medium to High	High	High	High	Low	High
Cost[b]	$200	$200	$10k	$300	$150k	$150k	$50	$10k~$100k	$20k	$5k	$50	$400

[a] Because fruit types, sample sizes, and metrics vary, we coarsely classify accuracy into *high*, *medium*, and *low*, with systems marked as high accuracy if the NRMSE is <10% tolerance.
[b] Costs are estimated for the sensing hardware only; ancillary items (e.g. a PC or mobile phone) are excluded.

system completes a ripeness assessment in under 3.5 seconds, significantly faster than the commercial Felix F-750 (~10 seconds) [21]. Notably, the time can be further reduced through optimization, making the system suitable for high-traffic environments like supermarkets.

10 Related Work

The fruit industry and research community have tried different approaches for non-invasive fruit ripeness estimation. Table 3 summarizes the comparison between these systems.

10.1 RF-Based Approach

Researchers have explored various RF modalities for fruit sensing, including Wi-Fi, millimeter wave (mmWave), and sub-terahertz (sub-THz) signals.

Wi-Fi-based approach: Early work leveraged Wi-Fi channel state information (CSI) to estimate fruit ripeness by observing changes in dielectric properties [57]. However, due to the coarse resolution of Wi-Fi signals, these methods only offer qualitative assessments (e.g., unripe to overripe). Further, Wi-Fruit [33] combines Wi-Fi CSI with vision information to improve accuracy but depends on precise phase calibration before and after placing the fruit, making it unreliable in dynamic environments. In contrast, FruitScope uses received signal strength from FMCW radar and acoustic profiles, both of which are robust to dynamic environments.

MmWave-based approach: Researchers also use mmWave technology for fruit ripeness sensing. For example, [71] investigated the feasibility of estimating the sugar content of fruits using RSS of mmWave signals. However, their setup is costly and they need to test the fruit at multiple positions to improve accuracy, limiting the practicality. Another work used low-cost mmWave FMCW radar [58], but the low measurement granularity hinders the accuracy. In contrast, FruitScope addresses these limitations through the advanced frame design, integration of ultrasound modality, and a learning-based method that delivers accurate estimation from a single measurement position.

Sub-THz-based approach: Recent systems such as AgriTera [2] and Meta-Sticker [22] employ sub-terahertz (sub-THz) waves to assess fruit ripeness, with Meta-Sticker introducing a metamaterial tuned to the fruit's dielectric properties. While these solutions offer accurate sensing, they require significantly expensive hardware (>$150k), and in Meta-Sticker's case, per-fruit tagging, making them unsuitable for large-scale deployment. Even future sub-THz integration in 6G may lower costs, devices meeting the requirements of fruit sensing (wide bandwidth and high resolution) will likely remain more expensive than other RF technologies such as FMCW radars. In contrast, FruitScope eliminates the need for tags and uses low-cost devices for a more scalable and affordable solution.

10.2 Other Sensing Modalities

Vision-based approach: Traditional computer vision algorithms focus on detecting surface defects or color variations from standard fruit images [4, 26, 37], while recent research [20, 28] applied hyperspectral imaging techniques to obtain spectral signatures from a phone camera to determine ripeness stages. However, these vision-based methods cannot penetrate the peels and therefore are fundamentally constrained by their reliance on surface characteristics, as the peel properties often do not correlate with the internal properties of many fruit types.

Near-Infrared-based approach: Near-Infrared (NIR) spectroscopy evaluates fruit ripeness by analyzing light absorption patterns in the near-infrared range [51], with validated applications for avocados, mangoes, and pears [10, 14, 47]. However, accurate results require spectrometers with wide wavelength ranges and high spectral resolution [30], resulting in instruments costing $10k-$100k [13, 35] that are impractical for widespread adoption [33]. One approach uses a single green LED-photodiode pair to reduce the cost [78] but requires 90-second measurements to achieve low variance.

Vibration-based approach: Vibration-based methods assess fruit ripeness by analyzing internal mechanical properties [12]. Free vibration approaches observe natural resonance from air pulses or lasers [17, 19, 66] but require complex setups. Forced vibration methods use external acoustic stimulus. For example, [23] uses contact speakers and microphones to analyze sound transmission, though this yields poor accuracy and operates at 20 Hz-10 kHz frequencies that may cause discomfort. Ultrasound-based methods analyze signal attenuation or velocity at 50-100 kHz [36, 52, 65], but require 5-7 N contact force [5, 7, 24], risking fruit damage and increasing complexity and measurement time [53].

11 Discussion

In this section, we discuss the limitations of our design and propose potential solutions to enhance its capabilities.

Placement geometry. Our prototype fixes the fruit's placement in a tray to get rid of errors induced by the distance and other disturbances. This is a practical choice for retail and self-checkout kiosks, where the customers can place the fruit to the device and assess the fruit. We have emulated the placement variances during our data collection, and proposed a two-stage learning technique to achieve low STD among multiple measurements for a single specimen. On the other hand, our current setup is not a strict restriction: the tray can be replaced by a shallow rail or belt at factories, or by a height-adjustable shelf in self-checkout kiosks. Our label-efficient data collection pipeline, along with the learning algorithm, can help the system quickly adapt to new geometry.

Dataset size. We only evaluated four climacteric fruits (N = 80) to demonstrate generality across ripeness profiles. Larger cultivar diversity will further stress-test the model, and we believe expanding the dataset size could improve the performance; fortunately, our contrastive stage requires no labels, so collecting unlabeled rotations at packing houses is straightforward and inexpensive.

Generalizing to Diverse Produce. Our system is currently optimized for medium-sized fruits. Small items (<3 cm) lacks sufficient radar cross-section while large produce (>20 cm) exceeds current rig dimensions, presenting physical constraints. Nevertheless, these might be handled through batch estimation or mechanical scaling of the hardware. For produce with thick skins where signal attenuation is higher, the system provides a unique advantage. While dense rinds may dampen acoustic responses, mmWave radar remains promising for sensing surface-level dielectric shifts and moisture changes, as demonstrated in prior work [71] utilizing mmWave for thick-skinned citrus. This suggests a promising potential for adaptive modality weighting, where the system might dynamically prioritize the most reliable sensor data to ensure robustness across a wider variety of peel structures.

Collectively, these limits are engineering, not conceptual; they do not undercut our central contribution that coupling fine-grained FMCW sensing with pseudo-noise-enabled acoustic sensing yields label-efficient, contact-force-free quality estimation at <$400.

12 Conclusion

In this work, we present FruitScope, a multi-modal low-cost, easy-to-deploy non-invasive system that combines FMCW radar and acoustic sensors and is robust to environmental noises. In contrast to past work such as sub-THz systems, our system provides a cost-effective solution while maintaining high accuracy and practical deployability. To enable FruitScope, we introduced a novel FMCW frame structure, integrated with pseudo-noise-based acoustic sensing, and applied a label-efficient learning framework that achieves high accuracy in ripeness estimation with minimal labeled data. Beyond fruit sensing, we envision FruitScope enabling broader applications in precision agriculture, food processing, and distribution systems.

References

[1] 2023. How Many Fruits and Veggies Do We Waste Per Year? https://www.goruvi.com/blogs/news/644-million-tons-of-wasted-fruits-and-veggies-each-year. Accessed: 2025-11-06.

[2] Sayed Saad Afzal, Atsutse Kludze, Subhajit Karmakar, Ranveer Chandra, and Yasaman Ghasempour. 2023. AgriTera: Accurate Non-Invasive Fruit Ripeness Sensing via Sub-Terahertz Wireless Signals. In *Proceedings of the 29th Annual International Conference on Mobile Computing and Networking, ACM MobiCom 2023, Madrid, Spain, October 2-6, 2023*. ACM, 61:1–61:15. https://doi.org/10.1145/3570361.3613275

[3] Nicholas T Anderson, Phul P Subedi, and Kerry B Walsh. 2017. Manipulation of mango fruit dry matter content to improve eating quality. *Scientia Horticulturae* 226 (2017), 316–321.

[4] Israel Arzate-Vázquez, José Jorge Chanona-Pérez, María de Jesús Perea-Flores, Georgina Calderón-Domínguez, Marco A Moreno-Armendáriz, Hiram Calvo, Salvador Godoy-Calderón, Roberto Quevedo, and Gustavo Gutiérrez-López. 2011. Image processing applied to classification of avocado variety Hass (Persea americana Mill.) during the ripening process. *Food and Bioprocess Technology* 4, 7 (2011), 1307–1313.

[5] A. Bechar, A. Mizrach, P. Barreiro, and S. Landahl. 2005. Determination of Mealiness in Apples using Ultrasonic Measurements. *Biosystems Engineering* 91, 3 (2005), 329–334. https://doi.org/10.1016/j.biosystemseng.2005.04.008

[6] S. Boll. 1979. Suppression of acoustic noise in speech using spectral subtraction. *IEEE Transactions on Acoustics, Speech, and Signal Processing* 27, 2 (1979), 113–120. https://doi.org/10.1109/TASSP.1979.1163209

[7] F. Camarena and J.A. Martínez-Mora. 2006. Potential of ultrasound to evaluate turgidity and hydration of the orange peel. *Journal of Food Engineering* 75, 4 (2006), 503–507. https://doi.org/10.1016/j.jfoodeng.2005.04.052

[8] Changhao Chen, Stefano Rosa, Yishu Miao, Chris Xiaoxuan Lu, Wei Wu, Andrew Markham, and Niki Trigoni. 2019. Selective Sensor Fusion for Neural Visual-Inertial Odometry. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*.

[9] Ting Chen, Simon Kornblith, Mohammad Norouzi, and Geoffrey E. Hinton. 2020. A Simple Framework for Contrastive Learning of Visual Representations. In *Proceedings of the 37th International Conference on Machine Learning, ICML 2020, 13-18 July 2020, Virtual Event (Proceedings of Machine Learning Research, Vol. 119)*. PMLR, 1597–1607. http://proceedings.mlr.press/v119/chen20j.html

[10] CJ Clark, VA McGlone, C Requejo, A White, and AB Woolf. 2003. Dry matter determination in 'Hass' avocado by NIR spectroscopy. *Postharvest Biology and Technology* 29, 3 (2003), 301–308.

[11] Djork-Arné Clevert, Thomas Unterthiner, and Sepp Hochreiter. 2016. Fast and Accurate Deep Network Learning by Exponential Linear Units (ELUs). In *4th International Conference on Learning Representations, ICLR 2016, San Juan, Puerto Rico, May 2-4, 2016, Conference Track Proceedings*, Yoshua Bengio and Yann LeCun (Eds.). http://arxiv.org/abs/1511.07289

[12] Chengqiao Ding, Zhe Feng, Dachen Wang, Di Cui, and Weihao Li. 2021-03. Acoustic vibration technology: Toward a promising fruit quality detection method. *Comprehensive Reviews in Food Science and Food Safety* 20, 2 (2021-03).

[13] Felix Instruments. 2025. Food Science Instruments. https://felixinstruments.com/food-science-instruments/. Accessed: 2025-11-06.

[14] Alex Goke, Sara Serra, and Stefano Musacchi. 2018. Postharvest Dry Matter and Soluble Solids Content Prediction in d'Anjou and Bartlett Pear Using Near-infrared Spectroscopy. *HortScience horts* 53, 5 (2018), 669 – 680. https://doi.org/10.21273/HORTSCI12843-17

[15] Francesco Gringoli, Matthias Schulz, Jakob Link, and Matthias Hollick. 2019. Free Your CSI: A Channel State Information Extraction Platform For Modern Wi-Fi Chipsets. In *Proceedings of the 13th International Workshop on Wireless Network Testbeds, Experimental Evaluation and Characterization (WiNTECH '19)*. 21–28. https://doi.org/10.1145/3349623.3355477

[16] Heyu Guo, Shanmu Wang, Ruichun Ma, Shiqi Jiang, Yasaman Ghasempour, Omid Abari, Baining Guo, and Lili Qiu. 2026. OmniVLA: Physically-Grounded Multimodal VLA with Unified Multi-Sensor Perception for Robotic Manipulation. In *2026 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE.

[17] Sam Hitchman, Kasper van Wijk, and Zoe Davidson. 2016. Monitoring attenuation and the elastic properties of an apple with laser ultrasound. *Postharvest Biology and Technology* 121 (2016), 71–77.

[18] Chiori Hori, Takaaki Hori, Teng-Yok Lee, Ziming Zhang, Bret Harsham, John R. Hershey, Tim K. Marks, and Kazuhiko Sumi. 2017. Attention-Based Multimodal Fusion for Video Description. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*.

[19] Naoki Hosoya, Michiru Mishima, Itsuro Kajiwara, and Shingo Maeda. 2017. Non-destructive firmness assessment of apples using a non-contact laser excitation system based on a laser-induced plasma shock wave. *Postharvest Biology and technology* 128 (2017), 11–17.

[20] Haiyan Hu, Yinan Zhu, Shanwen Chen, Qianyi Huang, and Qian Zhang. 2025. FruitPhone: Detecting Sugar Content in Fruits Using Unmodified Smartphones with Spectral Imaging. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 9, 3, Article 86 (Sept. 2025), 29 pages. https://doi.org/10.1145/3749470

[21] Felix Instruments. 2021. User Interface & Menu Walkthrough | F-750 + F-751. https://www.youtube.com/watch?v=RB-4B38Uf3E

[22] Subhajit Karmakar, Atsutse Kludze, and Yasaman Ghasempour. 2023. Meta-Sticker: Sub-Terahertz Metamaterial Stickers for Non-Invasive Mobile Food Sensing. In *Proceedings of the 21st ACM Conference on Embedded Networked Sensor Systems, SenSys 2023, Istanbul, Turkiye, November 12-17, 2023*. ACM, 335–348. https://doi.org/10.1145/3625687.3625815

[23] Hidetomo Kataoka, Takashi Ijiri, Jeremy White, and Akira Hirabayashi. 2016. Acoustic probing to estimate freshness of tomato. In *2016 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA)*. 1–5. https://doi.org/10.1109/APSIPA.2016.7820777

[24] Ki-Bok Kim, Sangdae Lee, Man-Soo Kim, and Byoung-Kwan Cho. 2009. Determination of apple firmness by nondestructive ultrasonic measurement. *Postharvest Biology and Technology* 52, 1 (2009), 44–48. https://doi.org/10.1016/j.postharvbio.2008.04.006

[25] Abhijit Kundu, Yin Li, and James M Rehg. 2018. 3d-rcnn: Instance-level 3d object reconstruction via render-and-compare. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 3559–3568.

[26] Shreya Lal, Santi Kumari Behera, Prabira Kumar Sethy, and Amiya Kumar Rath. 2017. Identification and counting of mature apple fruit based on BP feed forward neural network. In *2017 Third International Conference on Sensing, Signal Processing and Security (ICSSS)*. IEEE, 361–368.

[27] SK Lee, RE Young, PM Schiffman, and CW Coggins. 1983. Maturity studies of avocado fruit based on picking dates and dry weight. *J. Amer. Soc. Hort. Sci* 108, 3 (1983), 390–394.

[28] Zhiying Li1, Abdul Aziz, Patrick Phuoc, Phuc Nguye, and Tianxing Li. 2025. MobiChem: A Ubiquitous Smartphone-Based Toolkit for Practical Fruit Monitoring and Analysis. In *MobiSys '25: The 23rd Annual International Conference on Mobile Systems, Applications and Services, 2025*. ACM.

[29] Yumeng Liang, Anfu Zhou, Huanhuan Zhang, Xinzhe Wen, and Huadong Ma. 2021. FG-LiquID: A Contact-less Fine-grained Liquid Identifier by Pushing the Limits of Millimeter-wave Sensing. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 5, 3, Article 116 (Sept. 2021), 27 pages. https://doi.org/10.1145/3478075

[30] Hongjian Lin and Yibin Ying. 2009. Theory and application of near infrared spectroscopy in assessment of fruit quality: a review. *Sensing and instrumentation for food quality and safety* 3 (2009), 130–141.

[31] Yimeng Liu, Maolin Gan, Huaili Zeng, Li Liu, Younsuk Dong, and Zhichao Cao. 2024. Hydra: Accurate Multi-Modal Leaf Wetness Sensing with mm-Wave and Camera Fusion. In *Proceedings of the 30th Annual International Conference on Mobile Computing and Networking* (Washington D.C., DC, USA) *(ACM MobiCom '24)*. Association for Computing Machinery, New York, NY, USA, 800–814. https://doi.org/10.1145/3636534.3690662

[32] Yimeng Liu, Maolin Gan, Huaili Zeng, Yidong Ren, Gen Li, Jingkai Lin, Younsuk Dong, Xiaobo Tan, and Zhichao Cao. 2025. *Proteus: Enhanced mmWave Leaf Wetness Detection with Cross-Modality Knowledge Transfer*. Association for Computing Machinery, New York, NY, USA, 43–57. https://doi.org/10.1145/3715014.3722052

[33] Yutong Liu, Landu Jiang, Linghe Kong, Qiao Xiang, Xue Liu, and Guihai Chen. 2021. Wi-Fruit: See Through Fruits with Smart Devices. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 5, 4 (2021), 169:1–169:29. https://doi.org/10.1145/3494971

[34] Luna. 2018. The T-Ray® 5600/5700 Series Intelligent Terahertz Control Unit. https://lunainc.com/product/terahertz-control-unit-0 Accessed: 2025-11-06.

[35] Malvern Panalytical. 2024. ASD LabSpec 4 Hi-Res Analytical Instrument. https://www.malvernpanalytical.com/en/products/product-range/asd-range/labspec-range/labspec-4-hi-res-analytical-instrument. Accessed: 2025-11-06.

[36] Amos Mizrach. 2000. Determination of avocado and mango fruit properties by ultrasonic technique. *Ultrasonics* 38, 1 (2000), 717–722. https://doi.org/10.1016/S0041-624X(99)00154-7

[37] Vahid Mohammadi, Kamran Kheiralipour, and Mahdi Ghasemi-Varnamkhasti. 2015. Detecting maturity of persimmon fruit based on image processing technique. *Scientia Horticulturae* 184 (2015), 123–128.

[38] Fahim Niaz, Jian Zhang, Muhammad Khalid, Muhammad Younas, and Ashfaq Niaz. 2025. mmFruit: A Contactless and Non-Destructive Approach for Fine-Grained Fruit Moisture Sensing Using Millimeter-Wave Technology. *IEEE Transactions on Mobile Computing* 24, 5 (2025), 4022–4039. https://doi.org/10.1109/TMC.2024.3520914

[39] National Physical Laboratory (NPL). 1988. The Power Transmission Spectra of Possible Millimetre Wavelength Radome Materials. https://library.nrao.edu/public/memos/25/25M/25M_142.pdf

[40] Alan V Oppenheim and Ronald W Schafer. 2010. *Discrete-time Signal Processing*. Prentice Hall.

[41] Xiaomin Ouyang, Xian Shuai, Jiayu Zhou, Ivy Wang Shi, Zhiyuan Xie, Guoliang Xing, and Jianwei Huang. 2022. Cosmo: contrastive fusion learning with small data for multimodal human activity recognition. In *Proceedings of the 28th Annual International Conference on Mobile Computing And Networking*. 324–337.

[42] Malkeet S. Padda, Cassandro V.T. do Amarante, Raphael M. Garcia, David C. Slaughter, and Elizabeth J. Mitcham. 2011. Methods to analyze physico-chemical changes during mango ripening: A multivariate approach. *Postharvest Biology and Technology* 62, 3 (2011), 267–274. https://doi.org/10.1016/j.postharvbio.2011.06.002

[43] Markku Pukkila. 2000. Channel estimation modeling. *Nokia Research Center* 17 (2000), 66.

[44] M A Richards. 2022. *Fundamentals of radar signal processing*. Mcgraw Hill.

[45] James Rogers. 2021. See Beyond the Peel. https://www.apeel.com/blog/see-beyond-the-peel

[46] Margarita Ruiz-Altisent. 1991. Damage mechanisms in the handling of fruits. (1991).

[47] Sirinnapa Saranwong, Jinda Sornsrivichai, and Sumio Kawano. 2004. Prediction of ripe-stage eating quality of mango fruit from its harvest quality measured non-destructively by near infrared spectroscopy. *Postharvest Biology and Technology* 31, 2 (2004), 137–145. https://doi.org/10.1016/j.postharvbio.2003.08.007

[48] Manfred R Schroeder. 1979. Integrated-impulse method measuring sound decay without using impulses. *The Journal of the Acoustical Society of America* 66, 2 (1979), 497–500.

[49] Seeed Studio. 2025. ReSpeaker 6-Mic Circular Array Kit for Raspberry Pi. https://wiki.seeedstudio.com/ReSpeaker_6-Mic_Circular_Array_kit_for_Raspberry_Pi/. Accessed: 2025-11-06.

[50] Hailan Shanbhag, Sohrab Madani, Akhil Isanaka, Deepak Nair, Saurabh Gupta, and Haitham Hassanieh. 2023. Contactless Material Identification with Millimeter Wave Vibrometry. In *Proceedings of the 21st Annual International Conference on Mobile Systems, Applications and Services* (Helsinki, Finland) *(MobiSys '23)*. Association for Computing Machinery, New York, NY, USA, 475–488. https://doi.org/10.1145/3581791.3596850

[51] Syed Sohaib Ali Shah, Ayesha Zeb, Waqar S. Qureshi, Muhammad Arslan, Aman Ullah Malik, Waleed Alasmary, and Eisa Alanazi. 2020. Towards fruit maturity estimation using NIR spectroscopy. *Infrared Physics & Technology* 111 (2020), 103479. https://doi.org/10.1016/j.infrared.2020.103479

[52] Mahmoud Soltani Firouz, Ali Farahmandi, and Soleiman Hosseinpour. 2019. Recent advances in ultrasound application as a novel technique in analysis, processing and quality control of fruits, juices and dairy products industries: A review. *Ultrasonics Sonochemistry* 57 (2019), 73–88. https://doi.org/10.1016/j.ultsonch.2019.05.014

[53] Mahmoud Soltani Firouz, Ali Farahmandi, and Soleiman Hosseinpour. 2021. Early detection of freeze damage in navel orange fruit using nondestructive low intensity ultrasound coupled with machine learning. *Food Analytical Methods* 14 (2021), 1140–1149.

[54] Yonghao Song, Bingchuan Liu, Xiang Li, Nanlin Shi, Yijun Wang, and Xiaorong Gao. 2024. Decoding Natural Images from EEG for Object Recognition. In *The Twelfth International Conference on Learning Representations, ICLR 2024, Vienna, Austria, May 7-11, 2024*. OpenReview.net. https://openreview.net/forum?id=dhLIno8FmH

[55] Yonghao Song, Qingqing Zheng, Bingchuan Liu, and Xiaorong Gao. 2023. EEG Conformer: Convolutional Transformer for EEG Decoding and Visualization. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 31 (2023), 710–719. https://doi.org/10.1109/TNSRE.2022.3230250

[56] Yuqi Su, Fusang Zhang, Kai Niu, Tianben Wang, Beihong Jin, Zhi Wang, Yalan Jiang, Daqing Zhang, Lili Qiu, and Jie Xiong. 2024. Embracing Distributed Acoustic Sensing in Car Cabin for Children Presence Detection. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 8, 1, Article 16 (March 2024), 28 pages. https://doi.org/10.1145/3643548

[57] Sheng Tan, Linghan Zhang, and Jie Yang. 2018. Sensing Fruit Ripeness Using Wireless Signals. In *2018 27th International Conference on Computer Communication and Networks (ICCCN)*. 1–9. https://doi.org/10.1109/ICCCN.2018.8487344

[58] Reza Tavasoli, Sanjib Sur, and Srihari Nelakuditi. 2022. SSCense: a millimeter-wave sensing approach for estimating soluble sugar content of fruits. In *MobiSys '22: The 20th Annual International Conference on Mobile Systems, Applications and Services, Portland, Oregon, 27 June 2022 - 1 July 2022*. ACM, 553–554. https://doi.org/10.1145/3498361.3538780

[59] Texas Instruments. 2021. Introduction to mmWave Sensing: FMCW Radars. https://www.ti.com/lit/wp/spyy005a/spyy005a.pdf?ts=1719239577144. Accessed: 2025-11-06.

[60] Texas Instruments. 2021. mmWave Radar Radome Design Guide. https://www.ti.com/lit/an/swra705/swra705.pdf. Accessed: 2025-11-06.

[61] Texas Instruments. 2025. mmWave Sensing Estimator. https://dev.ti.com/gallery/view/mmwave/mmWaveSensingEstimator/ver/2.5.1/. Accessed: 2025-11-06.

[62] Texas Instruments. 2026. DCA1000EVM Evaluation Module. https://www.ti.com/tool/DCA1000EVM. Accessed: 2025-11-06.

[63] Shijie Tian and Huirong Xu. 2023. Mechanical-based and optical-based methods for nondestructive evaluation of fruit firmness. *Food Reviews International* 39, 7 (2023), 4009–4039.

[64] Chin-Chong Tseng and C Liu. 1972. Complementary sets of sequences. *IEEE Transactions on Information theory* 18, 5 (1972), 644–652.

[65] Marc Valente, Alexia Prades, and Didier Laux. 2013. Potential use of physical measurements including ultrasound for a better mango fruit quality characterization. *Journal of Food Engineering* 116, 1 (2013), 57–64. https://doi.org/10.1016/j.jfoodeng.2012.11.022

[66] Steve Vanlanduit, Frank Daerden, and Patrick Guillaume. 2007. Experimental modal testing using pressurized air excitation. *Journal of sound and vibration* 299, 1-2 (2007), 83–98.

[67] Saeed V Vaseghi. 2008. *Advanced digital signal processing and noise reduction*. John Wiley & Sons.

[68] Zhu Wang, Yifan Guo, Zhihui Ren, Wenchao Song, Zhuo Sun, Chao Chen, Bin Guo, and Zhiwen Yu. 2024. LiqDetector: Enabling Container-Independent Liquid Detection with mmWave Signals Based on a Dual-Reflection Model. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 7, 4, Article 186 (Jan. 2024), 24 pages. https://doi.org/10.1145/3631443

[69] Chenshu Wu, Feng Zhang, Beibei Wang, and K. J. Ray Liu. 2020. mSense: Towards Mobile Material Sensing with a Single Millimeter-Wave Radio. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 4, 3, Article 106 (Sept. 2020), 20 pages. https://doi.org/10.1145/3411822

[70] Xin Yang, Yunshuang Li, Yao Gong, Dinesh Jayaraman, Omid Abari, and Mingmin Zhao. 2026. Motion Capture with Millimeter-Wave Tags. In *Proceedings of the 24th ACM Conference on Embedded Networked Sensor Systems, SenSys 2026, Saint-Malo, France, May 11-14, 2026*. Association for Computing Machinery, New York, NY, USA.

[71] Zhicheng Yang, Parth H. Pathak, Mo Sha, Tingting Zhu, Junai Gan, Pengfei Hu, and Prasant Mohapatra. 2019. On the feasibility of estimating soluble sugar content using millimeter-wave. In *Proceedings of the International Conference on Internet of Things Design and Implementation, IoTDI 2019, Montreal, QC, Canada, April 15-18, 2019*. ACM, 13–24. https://doi.org/10.1145/3302505.3310065

[72] Hui-Shyong Yeo, Gergely Flamich, Patrick Schrempf, David Harris-Birtill, and Aaron Quigley. 2016. Radarcat: Radar categorization for input & interaction. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*. 833–841.

[73] Sangki Yun, Yi-Chao Chen, Huihuang Zheng, Lili Qiu, and Wenguang Mao. 2017. Strata: Fine-Grained Acoustic-based Device-Free Tracking. In *Proceedings of the 15th Annual International Conference on Mobile Systems, Applications, and Services, MobiSys'17, Niagara Falls, NY, USA, June 19-23, 2017*. ACM, 15–28. https://doi.org/10.1145/3081333.3081356

[74] Wen Zhang, Zhenzhen Lv, Bing Shi, Zihan Xu, and Lijun Zhang. 2021. Evaluation of quality changes and elasticity index of kiwifruit in shelf life by a nondestructive acoustic vibration method. *Postharvest Biology and Technology* 173 (2021), 111398. https://doi.org/10.1016/j.postharvbio.2020.111398

[75] Yi Zhang, Weiying Hou, Zheng Yang, and Chenshu Wu. 2023. VeCare: Statistical Acoustic Sensing for Automotive In-Cabin Monitoring. In *20th USENIX Symposium on Networked Systems Design and Implementation, NSDI 2023, Boston, MA, April 17-19, 2023*. USENIX Association, 1185–1200. https://www.usenix.org/conference/nsdi23/presentation/zhang-yi

[76] Flora Zidane, Jérôme Lanteri, Laurent Brochier, Nadine Joachimowicz, Hélène Roussel, and Claire Migliaccio. 2020. Damaged Apple Sorting With mmWave Imaging and Nonlinear Support Vector Machine. *IEEE Transactions on Antennas and Propagation* 68, 12 (2020), 8062–8071. https://doi.org/10.1109/TAP.2020.3016184

[77] Flora Zidane, Jerome Lanteri, Julien Marot, Laurant Brochier, Nadine Joachimowicz, Helene Roussel, and Claire Migliaccio. 2020. Nondestructive Control of Fruit Quality via Millimeter Waves and Classification Techniques: Investigations in the Automated Health Monitoring of Fruits. *IEEE Antennas and Propagation Magazine* 62, 5 (2020), 43–54. https://doi.org/10.1109/MAP.2020.3003222

[78] Agustin Zuniga, Huber Flores, and Petteri Nurmi. 2021. Ripe or Rotten? Low-Cost Produce Quality Estimation Using Reflective Green Light Sensing. *IEEE Pervasive Computing* 20, 3 (2021), 60–67. https://doi.org/10.1109/MPRV.2021.3074474