HearLiquid: Nonintrusive Liquid Fraud Detection Using Commodity Acoustic Devices

Yanni Yang, Yanwen Wang, *Member, IEEE*, Jiannong Cao, *Fellow, IEEE*, and Jinlin Chen

***Abstract*—Liquid fraud has plagued people with huge health risks. Liquid fraud detection can help to reduce the risk of liquid hazards. However, existing systems that use biochemical tools or radio frequency signals for liquid sensing are either expensive, intrusive, or inconvenient for public use. In this article, we propose HearLiquid, a low-cost and nonintrusive liquid fraud detection system using commodity acoustic devices. Our insight comes from the fact that acoustic impedance of different liquids results in distinct absorption of the acoustic signal across different frequencies when it travels through the liquid. In specific, we extract the liquid's acoustic absorption and transmission curve (AATC) over multiple frequencies of the acoustic signal for liquid fraud detection. However, accurately measuring the AATC faces multiple challenges. First, due to the hardware diversity and imperfection, different acoustic devices introduce diverse frequency responses, which brings significant deviations to AATCs of the same liquid. Second, different relative positions between acoustic devices and the liquid container result in variations in the AATC, making the detection result inaccurate. To overcome these challenges, we first calibrate the AATC using a dedicated reference AATC to remove the effect of hardware diversity. To bear the variations in AATCs measured from different relative positions, we apply a well-orchestrated data augmentation technique to automatically generate sufficient AATCs for different positions using a small number of collected data. Finally, AATCs are used to train the liquid detection model. We conduct extensive experiments on many important liquid fraud cases and achieve liquid detection accuracy of 92%–97%.**

***Index Terms*—Acoustic absorption and transmission, acoustic signal, liquid fraud detection.**

I. Introduction

LIQUID counterfeiting and adulteration have been jeopardizing human health for many decades. Fake liquids pose huge health risks and result in a large number of poisoning cases every year [1]. Common liquid fraud involves adulterating the expensive authentic liquid with cheaper and even harmful liquids or counterfeiting the authentic liquid with a similar flavor but different components. Adulterated and counterfeiting liquids are difficult to detect for consumers since they are camouflaged with the same appearance like the authentic one while only with the fake liquids inside. In recent years, governments, industries, and academia have taken great efforts to fight against liquid adulteration and counterfeiting [2], [3]. However, consumers still suffer from high risks owing to the lack of efficient and ubiquitous liquid detection approaches. This drives researchers to keep investigating better methods to detect liquid fraud.

Existing solutions for liquid detection can be mainly classified into four categories. The first category uses chemical and chromatographic techniques [4], [5]. These techniques enable precise detection of contaminants in the liquid. However, chemical tools and chromatographic equipment are quite cumbersome and expensive. For example, one set of infrared spectrometer could cost around U.S. \$15 000. Besides, chemical testings require direct contact with the liquid, which is intrusive for sealed liquids. The second category refers to the quasistatic electrical tomography (QET) technique [6], which measures the dielectric constant and conductivity of the liquid to detect flammable and explosive liquids in public areas. However, current QET systems only detect whether a liquid is flammable or explosive and are unable to detect liquid fraud. The third category measures the surface tension of the liquid to detect the liquid type using the tensiometer [7] or camera [8]. Nevertheless, surface tension measurement inevitably requires to open the liquid container. The fourth category leverages radio frequency (RF) signals, e.g., RFID [9], [10] and ultrawide band (UWB) radar[11], to measure liquid properties. The intuition is that RF signals traveling through or reflected by different liquids show different patterns of the signal parameter, e.g., the phase [12], [13] or Time-of-Flight (ToF) [11], which can be used for liquid detection. However, RF-based methods require specialized devices and cannot detect liquids with metal containers since metals could affect the normal transmission and communication of the RF signals, which limits its usage scenarios. Considering the limitations of existing liquid detection solutions, we ask a question: *can we detect liquid fraud in a cost-effective, nonintrusive, and ubiquitous manner?*

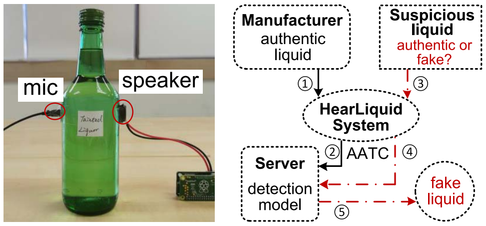
In this article, we propose HearLiquid, a liquid fraud detection system using commodity acoustic devices, i.e., the speaker and microphone. To the best of our knowledge, HearLiquid is the first to employ low-cost commodity acoustic devices for liquid fraud detection. In HearLiquid, the speaker and microphone are clung to the surface of the liquid container on each

Manuscript received September 11, 2021; revised December 8, 2021; accepted January 7, 2022. This work was supported in part by HK RGC Research Impact Fund under Grant R5034-18 and Grant R5060-19; in part by the National Nature Science Foundation of China under Grant 62102139; and in part by the Fundamental Research Funds for the Central Universities under Grant 531118010612. *(Corresponding author: Yanwen Wang.)*

Yanni Yang, Jiannong Cao, and Jinlin Chen are with the Department of Computing, The Hong Kong Polytechnic University, Hong Kong (e-mail: yan-ni.yang@connect.polyu.hk; jiannong.cao@polyu.edu.hk; csjlchen@comp.polyu.edu.hk).

Yanwen Wang is with the College of Electrical and Information Engineering, Hunan University, Changsha 410012, China (e-mail: wangyw@hnu.edu.cn).

Digital Object Identifier 10.1109/JIOT.2022.3144427



(a) (b)

Fig. 1. Demonstration of the HearLiquid system. (a) System setting. (b) Applied scenario.

side horizontally, as shown in Fig. 1(a). The speaker emits the acoustic signal, and the microphone receives the acoustic signal traveling through the liquid. Our key finding is that the received acoustic signal can be used to detect liquid fraud. The insight comes from the fact that liquids with different components have different acoustic impedance, which determines the absorption of the acoustic signal [14], [15]. Thus, the acoustic signal traveling through the liquid has the potential to distinguish the fake liquids from the authentic one by detecting the difference between the acoustic absorption patterns of the authentic and fake liquids. To measure the liquid's absorption of the acoustic signal, we extract the acoustic absorption and transmission curve (AATC) from the received acoustic signal. AATC characterizes the remaining energy of the acoustic signal over different frequencies after it travels through the liquid. In our work, the acoustic signal is dedicatedly generated and processed to extract the liquid's AATC, and we manage to remove the effects of several nonnegligible practical factors on AATC extraction.

Our system can detect whether a liquid is counterfeited or adulterated toward the authentic liquid product in real time, even without opening the liquid container. As shown in Fig. 1(b), authentic liquid manufacturers can use our system to extract liquid's AATCs, store them in the server, and train the liquid fraud detection model. When there is a suspicious fake, its AATC will be measured using the system and sent to the model stored in the manufacturer's server to get the detection result. We note that our system is not to replace existing liquid detection techniques (e.g., biochemical analysis and QET), but to provide a complementary technique that could allow consumers to detect liquid fraud in stores or at home.

Accurately extracting the AATC from the received acoustic signal is not a trivial task. The key challenges arise from the fact that AATC can be affected by many practical factors, which could lead to inaccurate liquid detection results. The first factor comes from the diversity and imperfection of the commodity acoustic devices. The frequency responses of different speakers and microphones vary a lot, even for acoustic devices from the same manufacturer. When extracting the liquid's AATC using different acoustic devices, the frequency response deviations result in inconsistent AATCs for the same liquid and consequently degrade the liquid detection accuracy. To tackle this problem, we propose a reference signal to remove the effect of acoustic devices' frequency responses. In specific, we place the speaker and microphone close to each other without space between them to measure a reference signal. The reference signal is mainly determined by the acoustic devices' frequency responses. Then, we extract the AATC of the acoustic signal traveling through the liquid. The liquid's AATC is decided by the liquid and most importantly, the same devices' frequency responses. By canceling out the acoustic devices' frequency responses using the reference signal, the effect of hardware diversity can be eliminated.

Another practical factor arises from the different relative device-container positions. In practice, the relative positions between the acoustic devices and liquid container may be hard to keep the same when using the system at different times. The position differences result in the change of multipath acoustic signals traveling inside the liquid container, which brings variations in the AATCs measured from the same liquid. Our experimental results show that the AATC variations could reduce the liquid detection accuracy. To address this problem, intuitively, we can extract AATCs from as many as possible positions to train the liquid detection model. However, it is labor intensive to collect such a large amount of training data. Thus, we adopt a data augmentation method to automatically generate AATCs for different relative device-container positions. Based on our observation that the AATCs measured from different positions follow the same distribution, we use the variational autoencoder (VAE) to generate AATCs for different positions using a small number of manually measured AATCs. However, due to the frequency-selective effect of acoustic signals, AATCs of the same liquid collected from different device-container positions exhibit a special variation pattern. As a result, multipath signals caused by the position difference could be strengthened or weakened on some frequencies. Based on this key observation, instead of directly applying existing VAE models, we improve the VAE by dedicatedly designing a *frequency-sensitive regularizer* in the original VAE loss function. Our AATC augmentation method can effectively improve the detection accuracy in the face of the effect from different device-container positions.

In this article, we make the following key contributions.

1) We propose HearLiquid, which, to the best of our knowledge, is the first work that uses commodity acoustic devices to detect liquid fraud. We extract a key feature from the acoustic signal traveling through the liquid, i.e., the AATC, for liquid fraud detection.
2) We perform an in-depth analysis of the practical factors that affect AATC extraction and tackle the corresponding challenges, including the effects of acoustic devices' frequency responses and different relative device-container positions, for accurate liquid fraud detection.
3) We implement the HearLiquid system and evaluate its performance with extensive experiments on various liquid fraud cases. The experimental results show that our system can achieve liquid fraud detection with an average accuracy of around 92%–97%.

II. Understanding the Acoustic Absorption and Transmission in Liquids

A. Liquid's Absorption of the Acoustic Signal

Our system employs the liquid's absorption of acoustic signal for liquid detection. The acoustic energy can be

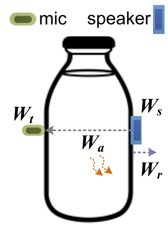


Fig. 2. Process of acoustic signal traveling through liquid.

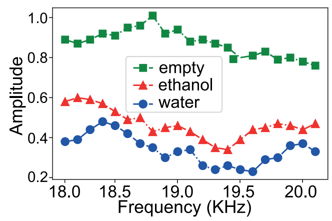


Fig. 3. Liquid absorption of acoustic signal in different medium.

absorbed when the acoustic signal travels through the liquid. This is because the acoustic pressure facilitates the movement of liquid particles, resulting in internal frictions caused by the viscosity effect, which converts the acoustic energy into heat and induces the absorption of acoustic signal [14]. The absorbed energy reaches its maximum when the acoustic frequency matches the liquid's natural frequency of vibration, i.e., the acoustic resonance phenomenon.[1] We model the process of transmitting the acoustic signal from the speaker on the right-hand side of the liquid to the microphone on the left-hand side in Fig. 2. During this process, the acoustic signal sent by the speaker (W_s) first encounters the liquid container. Then, part of the signal is reflected by the container surface (W_r). Part of the signal is absorbed by the liquid and transformed to heat (W_a). Finally, part of the signal travels through the liquid and is received by the microphone (W_t). If we keep the sent signal W_s and container unchanged, the energy of the received signal (W_t) is mainly decided by how much signal is absorbed in the liquid, i.e., W_a.

To show whether the acoustic signal can be affected by the liquid's absorption, we perform an experiment to compare the acoustic absorption without and with the water filled in a plastic bottle. We place one pair of speaker and microphone on two sides of the liquid container as shown in Fig. 2. Then, we transmit the acoustic signal with equal power on multiple frequencies, i.e., 18, 18.1, ..., 18.9, and 20 kHz. Then, we perform fast Fourier transformation (FFT) on the received signal and obtain the frequency-domain amplitude of each frequency. We keep all other settings unchanged during the experiment. As shown in Fig. 3, the amplitude of the empty bottle is higher than that of the bottle filled with water, showing that part of the sound energy is indeed absorbed by water.

The absorbed energy W_a is governed by the acoustic impedance (Z) of the liquid and is a function of frequency (f), i.e., $W_a(f) \sim Z_f$ [16]. The acoustic impedance is affected by the density of liquid (ρ) and the traveling speed (c) of acoustic signal in the medium, i.e., $Z = \rho \cdot c$ [17]. Since the density and sound speed are determined by liquid components, liquids with different components can result in different acoustic impedance. Thus, the absorbed energy W_a varies accordingly. To investigate the effect of different liquid components on the absorption of acoustic signal among multiple frequencies, we conduct an experiment to compare the acoustic absorption for two different liquids. We prepare water (density: 1.0 g/cc, speed: 1482 m/s under 25 °C) and ethanol (density: 0.79 g/cc, speed: 1159 m/s under 25 °C). Then, we fill the same amount of water and ethanol in the same containers and remain other settings unchanged. As shown in Fig. 3, the amplitudes of all frequencies for the received acoustic signal traveling through ethanol are larger than that through water, which shows that more energy is absorbed by water due to its higher density and sound speed than those of ethanol. In addition, amplitudes of the received signal vary among different frequencies across different liquids since W_a is affected by the sound frequency as well. Therefore, the AATC, which is composed of amplitudes over multiple frequencies of the acoustic signal after being absorbed and transmitting through the liquid, can serve as a good feature to differentiate different liquids. We will introduce the design of W_s, W_t processing, and AATC extraction in Section III.

B. Feasibility of Using AATC for Liquid Detection

To investigate the feasibility of using the AATC for distinguishing different liquids and liquid fraud detection, we first conduct a set of preliminary experiments to observe the AATCs for: 1) different kinds of liquids; 2) random mixtures of one liquid with other fraudulent liquids; and 3) mixtures of one liquid with different percentages of another fraudulent liquid. For 1), we select three kinds of alcohol products (liquor, ethanol, and isopropanol) and three kinds of cooking oil products (olive oil, canola oil, and soybean oil) in the market. For 2), we regard the ethanol and isopropanol as the fraudulent alcohol against the liquor and treat the canola oil and soybean oil as the fraudulent oil against the olive oil. Then, we randomly mix the liquor with ethanol and isopropanol, as well as mixing the olive oil with canola oil and soybean oil, respectively.

For 3), we mix the liquor and olive oil with different percentages of isopropanol (30% and 40%) and canola oil (20% and 30%), respectively. For each of the original and mixed liquids, we collect three traces of acoustic signal and extract the AATC from each trace. During the experiment, we use the same acoustic devices and container for all the liquids. We emit the acoustic signal with 21 frequencies ranging from 18 to 20 kHz with an interval of 100 Hz. Fig. 4(a)–(c) depicts the AATCs of the liquor, ethanol, isopropanol, and their mixtures.[2] The AATCs of the olive oil, canola oil, soybean oil, and their mixtures are shown in Fig. 4(d)–(f). From Fig. 4, we have the following observations.

1) The AATCs of the same liquid exhibit similar patterns and are stable across different times of measurements.

[1] The acoustic resonance phenomenon will rarely happen for our case since the resonant frequencies of liquids are around GHz-level, while the sound we transmit is in the 18–20-kHz frequency band.

[2] "etha" and "isop" in Fig. 4(b) and (c) are the abbreviations of the ethanol and isopropanol, respectively. (r) refers to the random mixture.

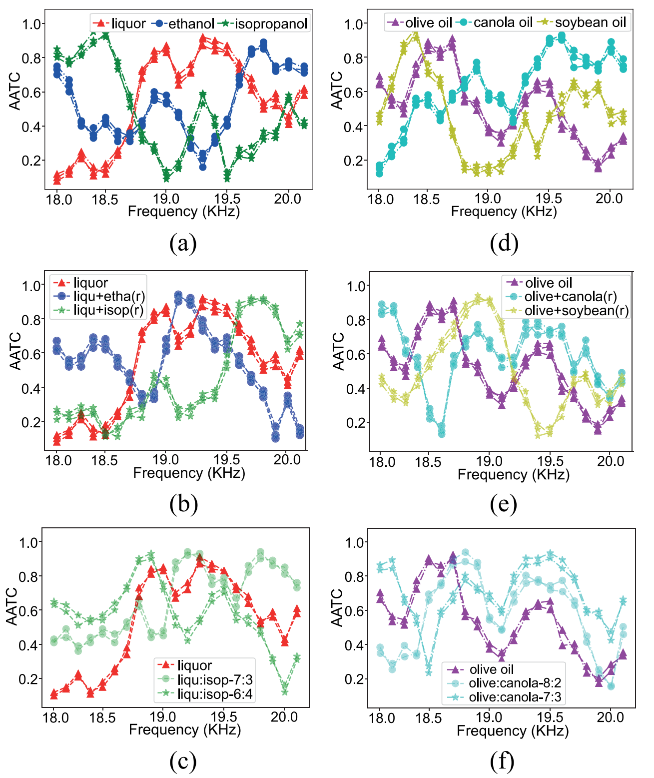


Fig. 4. AATCs of different authentic and adulterated liquids. (a) AATCs of liquor, ethanol, and isopropanol. (b) AATCs of liquor mixed with etha and isop. (c) AATCs of different % of isop mixed liquor. (d) AATCs of olive, canola, and soybean oil. (e) AATCs of olive mixed with canola and soybean. (f) AATCs of different % of canola mixed olive.

2) For different kinds of liquids, as shown in Fig. 4(a) and (d), the AATCs show distinct patterns, indicating that the AATC is potential to distinguish different liquids.
3) As shown in Fig. 4(b) and (e), the AATCs of the authentic liquids are distinct from those of the fake liquids mixed with different fraudulent liquids. Furthermore, as depicted in Fig. 4(c) and (f), the AATCs of the authentic liquids compared with those of the fake liquids mixed with different percentages of the fraudulent liquid are different as well. This indicates that it is potential to use the AATC for detecting the fake liquids with different kinds and percentages of fraudulent liquids out of the authentic liquid.

We also build a simple anomaly detection model using AATCs of the authentic liquid to obtain preliminary liquid detection results. We collect 125 AATCs from the authentic liquor and olive oil, respectively. Seventy five AATCs are used to train a one-class support vector machine (SVM) model for anomaly detection, and the left 50 AATCs are used for testing. We also collect 50 AATCs from each of the fraudulent and fake liquids. The accuracy for fake liquor and olive oil detection reaches 86.5% and 84.7%, respectively. Our observations and experimental results show that it is feasible to use AATC for liquid fraud detection.

C. Practical Factors for AATC Extraction

Although AATC is useful for liquid fraud detection, in practice, AATC extraction is vulnerable to multiple practical factors, which may significantly affect the liquid detection result. The first factor comes from the hardware diversity and imperfection of acoustic devices. The frequency responses of commodity speakers and microphones vary a lot across different frequencies, especially for the high-frequency band above 17 kHz [18]. The acoustic devices' frequency responses could affect the frequency-domain amplitudes of the received acoustic signal, resulting in inconsistent AATCs for the same liquid. To show the effect of different acoustic devices on AATC extraction, we use the same liquid but apply two different pairs of speakers and microphones to send and receive the signal. In Fig. 5(a), the AATCs exhibit dissimilar patterns when using different acoustic devices for the same liquid. Hence, the AATC can be greatly affected by the frequency responses of the acoustic devices. In Section III-D, we will introduce our proposed AATC calibration method to remove the effect of acoustic devices' frequency responses.

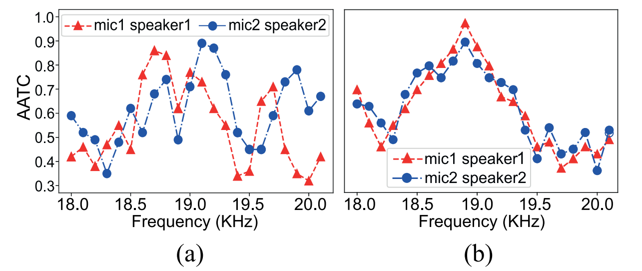


Fig. 5. Raw and calibrated AATCs of the same liquid using different speakers and microphones. (a) AATCs of different acoustic devices. (b) Calibrated AATCs in (a).

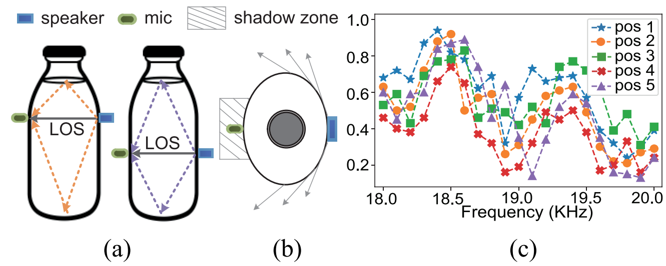


Fig. 6. Effect of relative device-container positions. (a) Signals in liquid. (b) Acoustic shadow zone. (c) AATCs of different positions.

The second factor lies in the different relative positions between acoustic devices and liquid container. In practice, the position of acoustic devices relative to the container may be hard to keep the same when using the system at different times. The change of the relative device-container position can affect the AATC. This is because the acoustic signal traveling from different parts of the container can result in different multipath signals inside the container. Fig. 6(a) shows the propagation paths of the acoustic signal in the liquid. The Line-of-Sight (LoS) signal remains unchanged when the acoustic devices are placed at different heights relative to the container. However, multipath signals reflected by the liquid and container (dashed lines) change along with different positions. Those changes result in variations of the received acoustic signal, making the AATCs measured from different positions vary for the same liquid. To show the effect of different positions on the AATC, we place the same acoustic devices at five different heights of

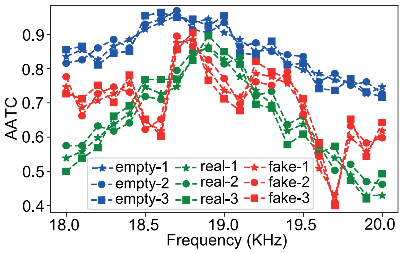


Fig. 7. AATCs for three same bottles when empty, filled with authentic, and fake wine.

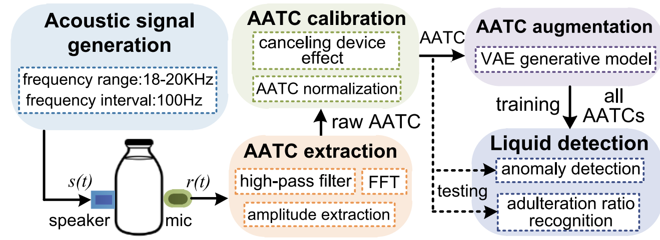


Fig. 8. Overview of the HearLiquid system.

the container. Then, we extract the AATCs for the olive oil at these five positions, as shown in Fig. 6(c). The AATCs for the same liquid exhibit variations across the frequency band under different relative positions. These variations in the AATC could lead to misdetection of the same liquid. We perform preliminary experiments on detecting the authentic and fake olive oil with the same acoustic devices placed at different relative positions to the container. We use the authentic olive oil's AATCs collected at one position to train the one-class SVM model. Then, the AATCs of the authentic and fake olive oil collected at another four positions are used for testing. The detection accuracy decreases to 71.3% with errors mainly coming from inaccurately detecting the authentic olive oil as the fake one. An intuitive solution to promote the accuracy is to collect the acoustic signal from as many as relative device-container positions for the authentic liquid to train the anomaly detection model. However, it is labor intensive or even impractical to collect such a large number of data. Thus, an alternative method is needed to deal with the insufficient training data. We will elaborate on our data augmentation method in Section III-E.

There are other factors that affect the AATC, including the container, incident angle of the acoustic signal, sound diffraction, temperature, and humidity. First, the liquid container could affect the AATC because part of the sound energy is inevitably absorbed by containers. While after fully considering the liquid fraud in practice that most fake liquids are filled into the same container and package as the authentic one to deceive consumers so that people cannot detect them by the appearance, we can have a reasonable assumption that the containers' effect can be regarded as an identical constant for authentic and fake liquids. To investigate this assumption, we conduct experiments to compare AATCs for the same containers. First, we extract AATCs for three same plastic bottles filled with the same amount of authentic wine. Then, we mix all three bottles of authentic wine with the same amount of cheap ethanol as the fake wine and extract their AATCs. Finally, we empty the three bottles and extract their AATCs. As shown in Fig. 7, AATCs for the three bottles all share similar patterns when empty, filled with authentic wine, and filled with fake wine, respectively. This experiment result indicates that the container effect can be regarded as a constant factor for the same containers and can be neglected during data collection.

Second, the acoustic absorption can be affected by the acoustic signal's incident angle [19]. This effect can be avoided by our system setting, in which the acoustic devices cling to the container's surface horizontally so that the incident angle is fixed. Third, due to the sound diffraction effect, the acoustic signal may bypass the container and continue to travel behind it. The diffracted signal may be superimposed with the signal traveling through the liquid. However, the signal diffraction effect can be ignored if the acoustic signal's wavelength is smaller than the container [20]. In our case, the acoustic signal's wavelength (around 1.8 cm) is much smaller than the size of the liquid container (radius: 5–10 cm). Besides, the diffraction effect can be further mitigated in our system because the microphone is deployed in the acoustic shadow zone [21], as shown in Fig. 6(b). The diffracted signal is significantly reduced, and the microphone mainly receives the signal traveling through the liquid. Finally, the environmental temperature and humidity can affect the acoustic absorption of the liquid in a linear way [22], while normalizing the AATC can help to alleviate such impacts. Hence, our work mainly focuses on eliminating the effects of the diversity of acoustic devices and different relative device-container positions.

III. System Design

A. System Overview

The overview of the HearLiquid system is shown in Fig. 8. First, we generate the acoustic signal $s(t)$ which is emitted by the speaker. Then, the acoustic signal travels through the liquid and is received by the microphone as $r(t)$. Next, $r(t)$ is preprocessed and the raw AATC is extracted. Since the raw AATC includes the effect of the acoustic devices' frequency responses, we calibrate the AATC using a designated reference signal. In addition, to obtain AATCs for as many as different relative device-container positions in an effective way, we automatically augment the AATCs using a generative VAE model. Finally, AATCs are used to train the liquid detection models. For liquid fraud detection, an anomaly detection model is built using the AATCs of the authentic liquid. We further extend the functionality of HearLiquid to adulteration ratio recognition for fake liquids. To this end, we build a classification model using the AATCs of all the adulterated liquids to recognize the adulteration ratio. For an unknown liquid, we can use the anomaly detection model to detect whether the liquid is authentic or fake. Besides, the classification model can be applied to recognize the adulteration ratio for a fake liquid of interest.

B. Acoustic Signal Generation

The emitted acoustic signal $s(t)$ is designed as the sum of multiple sine waves with different frequencies, i.e., $s(t) = \sum_{i=1}^{n} A_i \sin(2\pi f_i t)$, where A_i is the amplitude of each sine

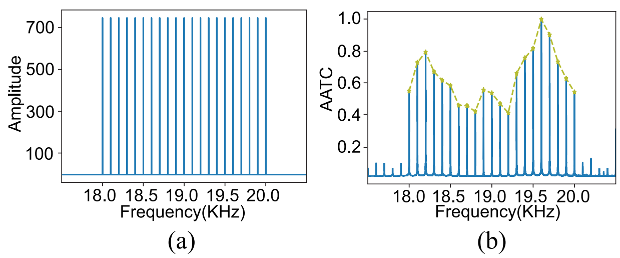


Fig. 9. Spectrum of generated acoustic signal and extracted AATC. (a) Spectrum of $s(t)$. (b) AATC of received signal.

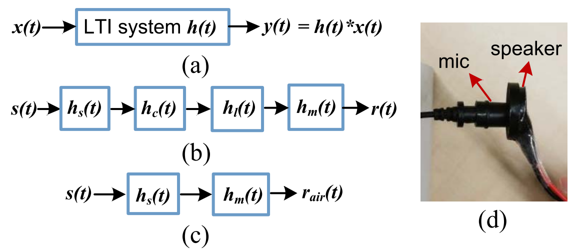


Fig. 10. Modeling of acoustic signal transmitting from speaker to microphone and the setting for AATC calibration. (a) Input, impulse response, and output of the LTI system. (b) Signal from the speaker, container, liquid, to mic. (c) Signal from speaker, the air, to mic. (d) Reference setting.

wave, f_i is the frequency, and n is the number of discrete frequencies. In our design, A_i is the same for all the sine waves. The discrete frequency f_i is within the frequency band of [18 kHz, 20 kHz]. The reason for choosing the frequency band of [18 kHz, 20 kHz] lies in four aspects. First, the acoustic signal in this frequency band is inaudible to most people, which does not disturb the users when using the system. Second, frequencies of most background noises in the environment, as well as the human voice, are lower than 18 kHz [23]. Then, the noises in the environment are removed with a high-pass filter. Third, the acoustic signal's wavelength within such a frequency band is much smaller than the size of most containers, which can alleviate the sound diffraction effect. Finally, the upper bound of the frequency for most commodity speakers and microphones is 20 kHz. The interval I_f between every two discrete frequencies is equal, and I_f determines the granularity of AATC. In Section IV-D2, we will discuss the effect of AATC granularity on the liquid detection performance. Finally, $s(t)$ is saved as a WAV file, which is played by the speaker. Fig. 9(a) shows the spectrum of $s(t)$ with 21 frequencies, i.e., $I_f = 100$ Hz.

C. Signal Preprocessing and AATC Extraction

After emitting the acoustic signal from the speaker, the microphone receives the acoustic signal for 4 s with a sampling rate of 48 kHz. Then, a high-pass filter with a cutoff frequency of 18 kHz is applied on the received acoustic signal $r(t)$ to remove the background noises. Next, we perform FFT on the filtered signals. A Hamming window is applied on the filtered signal before FFT to reduce the frequency leakage. Then, we extract the frequency-domain amplitude at f_i as $R(f_i)$. Finally, $R(f_i)$ is divided by the corresponding amplitude $S(f_i)$ in the spectrum of $s(t)$ to obtain the AATC. AATC represents the ratio of the remaining acoustic signal's energy over the emitted acoustic signal's energy across multiple frequencies. In practice, the volume of the speaker and microphone may change. Thus, we normalize the AATC to the same scale of [0, 1] after AATC calibration. Fig. 9(b) shows the normalized AATC, as denoted by the yellow-dashed curve.

D. AATC Calibration: Tackling the Effect of Different Acoustic Devices

1) Modeling the Transmission of the Acoustic Signal From the Speaker, Liquid, and Its Container to the Microphone: The transmission of the acoustic signal in the whole system can be modeled as follows.

1) *Frequency Responses of Speaker and Microphone:* The speaker and microphone are typical linear time-invariant (LTI) systems [24], which produce an output signal $y(t)$ from any input signal $x(t)$ subject to the constraints of linearity and time invariance. The characteristic of an LTI system is described by its impulse response $h(t)$. Fig. 10(a) shows the relationship among $x(t)$, $h(t)$, and $y(t)$ of the LTI system. The impulse responses of the speaker and microphone are denoted as $h_s(t)$ and $h_m(t)$, respectively, and the corresponding frequency responses are $H_s(f)$ and $H_m(f)$.
2) *Acoustic Signal Transmission in the Liquid and Its Container:* When the acoustic signal travels through a medium, due to the reflection of the obstacles in the medium, there are multiple paths of the signal with different delays arriving at the receiver. The received signal can be modeled as an LTI system as well [25], which can be expressed as $y(t) = \sum_{i=1}^{N} a_i x(t-\tau_i) = h(t) * x(t)$, where $h(t)$ is the signal's channel impulse response in the medium, N is the number of paths, and a_i and τ_i are the amplitude and time delay of each signal path, respectively. When the acoustic signal travels through the container, its channel impulse response $h_c(t)$ is mainly affected by the container's material and thickness. For the acoustic signal traveling through the liquid, i.e., $y_l(t) = \sum_{i=1}^{N_l} a_{l_i} x(t - \tau_{l_i}) = h_l(t) * x(t)$, a_{l_i} and τ_{l_i} in its channel frequency response $h_l(t)$ contain the information about the liquid's absorption of the acoustic signal.

Finally, as modeled in Fig. 10(b), the overall received acoustic signal $r(t)$ after the emitted signal $s(t)$ traveling through the cascade of the above four LTI systems can be expressed as $r(t) = s(t) * h_s(t) * h_c(t) * h_l(t) * h_m(t)$. By transforming $r(t)$ into the frequency domain, it becomes $R(f) = S(f) \cdot H_s(f) \cdot H_c(f) \cdot H_l(f) \cdot H_m(f)$, where $H_c(f)$ and $H_l(f)$ are the channel frequency responses in the container and liquid, respectively. When using different acoustic devices to measure the AATC for the same liquid and container, $S(f)$, $H_c(f)$, and $H_l(f)$ keep unchanged, while $H_s(f)$ and $H_m(t)$ are different, resulting in different $R(f)$ with inconsistent AATCs.

2) Calibrating the AATC Using the Reference Signal: To remove the effect of the acoustic devices' frequency responses, we design a reference signal, which directly travels from the speaker and microphone without other medium between them. Specifically, at the initialization stage of the system, we place the speaker and microphone close to each other without space

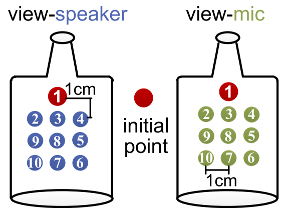


Fig. 11. Positions of speaker and microphone.

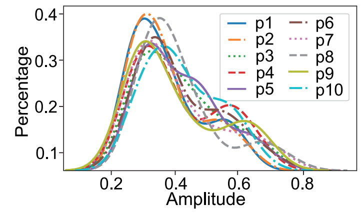


Fig. 12. Distribution of the same liquid's AATCs for devices at ten positions.

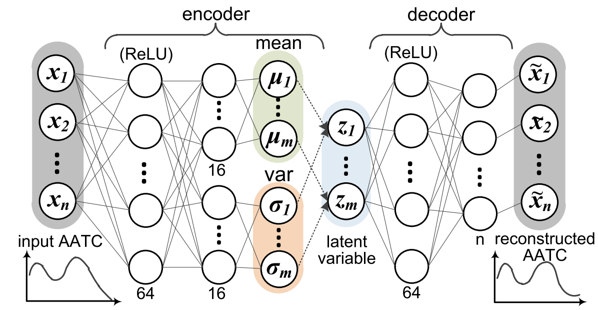


Fig. 13. Architecture of the VAE model which augments AATCs for different relative device-container positions.

between them, as shown in Fig. 10(d). Under this setting, the acoustic signal would directly travel from the speaker and microphone without other medium between them. Although there is still some air inside devices, the portion is quite small, so it can be neglected. Then, the measured reference signal $r_{\text{ref}}(t) = s(t) * h_s(t) * h_m(t)$. The corresponding frequency-domain representation becomes $R_{\text{ref}}(f) = S(f) \cdot H_s(f) \cdot H_m(f)$. Since we use the same acoustic devices to measure the reference signal and liquid's AATCs, dividing the reference signal by the signal traveling through the liquid in the frequency domain becomes the following equation:

$$\begin{aligned}\frac{R_{\text{ref}}(f)}{R_l(f)} &= \frac{S(f) \cdot H_s(f) \cdot H_m(f)}{S(f) \cdot H_s(f) \cdot H_c(f) \cdot H_l(f) \cdot H_m(f)} \\ &= \frac{1}{H_c(f)} \cdot H_l(f). \qquad (1)\end{aligned}$$

It shows that the calibrated signal is irrelevant to $H_s(f)$ and $H_m(f)$. In addition, $H_c(f)$ is a constant factor for the same type of liquids to be detected. Thus, the calibrated AATC is only affected by the liquid's frequency response $H_l(f)$, which can reflect the acoustic absorption of the liquid. Note that such a setting for calibration is a one-time setup before liquid detection, and the frequency response does not need to be calibrated again with the same speaker–microphone pair. Based on the above AATC calibration method, we calibrate the raw AATCs measured with different acoustic devices in Fig. 5(a). The calibrated AATCs are shown in Fig. 5(b). The calibrated AATCs when using different speakers or microphones exhibit similar patterns for the same liquid, which shows the effectiveness of our calibration method.

E. Data Augmentation: Tackling the Effect of Different Relative Device-Container Positions

Ideally, the collected AATCs should involve all the variations caused by different relative device-container positions to train the liquid detection model. However, manually collecting the AATCs from as many as possible positions is quite labor intensive. In our work, we adopt a data augmentation technique, which can automatically emulate the variations in AATCs caused by different relative device-container positions.

To find a proper method for AATC augmentation, we investigate the characteristics of the AATCs extracted from different relative device-container positions. We first choose one initial position at the center of the liquid container to place the speaker and microphone. Then, we move the speaker–microphone pair up, down, left, and right with 1-cm stepwise, as shown in Fig. 11. In sum, ten different pairs of positions are selected for the speaker and microphone. In Fig. 12, we depict the distribution of the AATCs extracted with the acoustic devices placed at ten different positions relative to the container for the same liquid. For each position, five AATCs are extracted. Fig. 12 shows that the extracted AATCs share similar distributions at different positions. We also measure the AATC's distribution for another ten liquids and observe similar patterns. We further apply the equivalence test on the AATCs of different positions to check whether they follow the same distribution. The equivalence interval is set to the average difference among the AATCs collected from the same position, i.e., 0.03 in our experiment. The average p-value is 0.019 (threshold as 0.05), which rejects the hypothesis that the difference among the AATCs of different positions is larger than the equivalence interval. This indicates that the same liquid shares the same AATC distribution even at different relative device-container positions. As such, we can employ the generative model, which can generate new data following the same distribution of the input data with some variations, to augment the AATCs. In our work, we employ VAE for AATC augmentation because it can effectively augment more data based on a small amount of input data [9], [26].

Fig. 13 shows the VAE model for AATC augmentation. The input $x(n)$ is the vector of AATC, which is extracted from the manually collected acoustic signal. The output $\widetilde{x}(n)$ is the reconstructed AATC. The VAE model consists of an encoder whose target is to compress the input feature vector into a latent variable vector $z(m)$ and a decoder that decompresses $z(m)$ to reconstruct the input. m is the length of the latent variable vector. Since the latent variable vector learns a representation with fewer dimensions than the input, m should be smaller than n. For our case with $n = 21$, $m < 21$. Meanwhile, considering that too few dimensions of z could lead to larger information loss, we empirically select $m = 16$, which, in our work, achieves the highest accuracy when using the generated AATCs to train the model for liquid fraud detection.

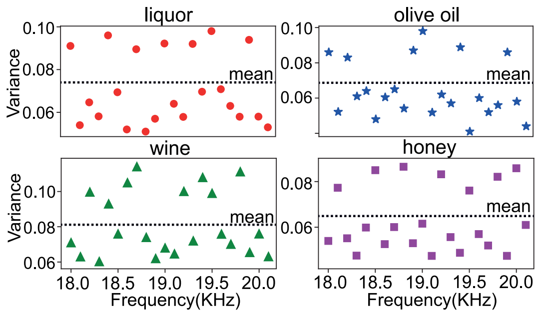


Fig. 14. Variances of the AATCs at different frequencies for different relative device-container positions.

To further enhance the performance of VAE for AATC augmentation, we add a regularizer in VAE's loss function based on a key observation about the AATCs of different relative device-container positions. We find that the AATCs on some frequencies experience larger variance than other frequencies at different positions. This is mainly due to the frequency-selective fading effect of the acoustic signal [27]. In specific, the change of multipath acoustic signals caused by the position difference could strengthen or weaken the amplitude of the received acoustic signal with a larger degree on some frequencies. To show the frequency-selective fading effect on the AATCs of different positions, we depict the variances of AATCs over all the frequencies for the authentic liquor, olive oil, wine, and honey in Fig. 14. It shows that AATCs exhibit larger variances at several frequencies, i.e., the frequencies whose variances are above the mean of all the variances. This indicates that some frequencies are more sensitive to different positions. Recall that our purpose of using VAE is to generate AATCs that seem like being obtained under different device-container positions. To this end, we add a *frequency-sensitive regularizer* in the VAE's loss function $\mathcal{L}$ to enlarge the AATCs' variances for those sensitive frequencies as follows:

$$\mathcal{L} = \underbrace{E_{q_{\theta(z|x_i)}}\left[\log p_{\phi}(x|z)\right]}_{\text{reconstruction loss}} - \underbrace{KL(q_{\theta}(z|x)\|p(z))}_{\text{KL divergence}}$$
$$- \underbrace{\left\|\mathrm{AATC}(f_{\mathrm{sen}}) - \tilde{\mathrm{AATC}}(f_{\mathrm{sen}})\right\|_1}_{\text{frequency-sensitive regularizer}}. \tag{2}$$

In (2), the first and second terms, which are the reconstruction loss between the input and generated AATCs and the Kullback–Leibler (KL) divergence, form the original VAE loss function. The third term is our added *frequency-sensitive regularizer*, where $\mathrm{AATC}(f_{\mathrm{sel}})$ and $\tilde{\mathrm{AATC}}(f_{\mathrm{sel}})$ are the input and generated AATCs' values on the sensitive frequencies f_{sen}, respectively. To select f_{sen}, we first calculate the variances of the manually measured AATCs for each frequency and obtain the mean of all the variances. Then, the frequencies whose variances exceed the mean are selected as f_{sen}. When training the VAE model, $\mathcal{L}$ is minimized to find the optimal weights in (2); meanwhile, the difference between the input and generated AATCs' values on those sensitive frequencies is enlarged. Finally, based on a certain number of manually measured AATCs, the VAE model will generate more AATCs for the authentic liquid, which are combined with the manually measured AATCs to train the liquid detection model.

F. Liquid Detection

1) Liquid Fraud Detection: To detect liquid fraud, intuitively, we can collect data from both authentic and fake liquids to train a binary classification model. However, in practice, it is difficult or impractical to acquire all kinds of fake liquids with various fraudulent components and adulteration ratios. Thus, we regard fake liquids as anomalies toward the authentic liquid and propose to build an anomaly detection model only using the AATCs of the authentic liquid.

We employ the VAE to build the anomaly detection model [28], [29]. The principle of using VAE for anomaly detection lies in the differences between the reconstruction losses of authentic and fake liquids. When training the VAE model using the authentic liquid's AATCs, the reconstruction loss between the input and generated AATCs is minimized. Then, VAE can learn how to generate new AATCs following the same distribution of the authentic liquid. When the testing input is the AATC of an authentic liquid, the reconstruction loss can be quite small. While if the testing input comes from a fake liquid since the AATCs of the authentic and fake liquids have different distributions, the reconstruction loss would be larger than that of the authentic liquid. Therefore, we can use the reconstruction loss of VAE to train the anomaly detection model. Specifically, we obtain all the reconstruction loss values when using the authentic liquid's AATCs to train the VAE model in Fig. 13. Then, we follow the three-sigma rule of thumb to select the threshold δ_t to detect the anomalies [30]. The mean (μ_t) and standard deviation (σ_t) of all the losses are calculated. We compare the liquid fraud detection accuracy using $\mu_t + \sigma_t$, $\mu_t + 2\sigma_t$, and $\mu_t + 3\sigma_t$ as δ_t, respectively. We set δ_t to $\mu_t + \sigma_t$ since it achieves the best accuracy. For an unknown liquid, we input its AATC to the VAE model. If the reconstruction loss is larger than δ_t, it is detected as the fake liquid, and *vice versa*.

2) Liquid Adulteration Ratio Recognition: Apart from liquid fraud detection, we find that the AATC has the potential to differentiate the adulterated liquids with different adulteration ratios according to the observations from Fig. 4(c) and (f). For instance, the AATCs of mixing the liquor with isopropanol by the ratios of 7:3 and 6:4 show different patterns. This brings the opportunity for recognizing the liquid adulteration ratio. In practice, the liquids could be harmful to human health if the adulteration ratio exceeds a certain level. Therefore, it would be useful to recognize the liquid adulteration ratio using the AATC.

To achieve this, we first predefine the interested adulteration ratios, e.g., 20%, 30%, and 40%, and collect the acoustic signal traveling through the adulterated liquids with different adulteration ratios. Next, AATC extraction, calibration, and augmentation are performed. Before inputting AATCs for training, we apply the largest margin nearest neighbor (LMNN) to map the AATC into a new space, so that the AATCs of the liquids with different adulteration ratios become more discriminative from each other. This is because LMNN

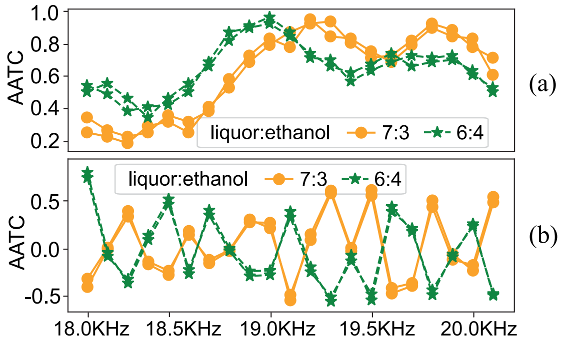


Fig. 15. AATCs before and after performing LMNN. (a) Before LMNN. (b) After LMNN.

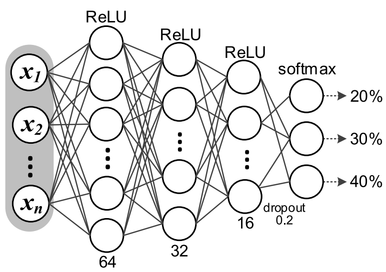


Fig. 16. MLP-based classification model for adulteration ratio recognition.

TABLE I
SPECIFICATION OF THE ACOUSTIC DEVICES

Device	Brand	Parameter	Frequency range
Speaker	iLouder	3W, 8Ω	700Hz - 20KHz
	iLouder	2W, 8Ω	700Hz - 20KHz
	HNDZ	3W, 4Ω	100Hz - 20KHz
Mic	Sony D11	/	65Hz - 20KHz
	BoYa M1	/	200Hz - 20KHz
	Dayton imm-6	/	80Hz - 20KHz

can "pull" the AATCs of the same class closer and meanwhile, "push" the AATCs of different classes farther from each other. By doing this, LMNN can find a space in which AATCs of different adulteration ratios become larger while AATCs of the same ratio are narrowed. Fig. 15 depicts the AATC before and after applying the LMNN for two ethanol-mixed liquor with close adulteration ratios. Compared with AATCs without LMNN, the transformed AATCs of the same liquid after LMNN are closer to each other, and AATCs of different liquids are of larger difference with each other. Finally, we apply the multilayer perceptron (MLP) neural network to build the classification model, as shown in Fig. 16. For the fake liquid with an unknown adulteration ratio, its AATC is extracted and input to the classification model to obtain the adulteration ratio.

IV. IMPLEMENTATION AND EVALUATION

A. Hardware

HearLiquid is implemented with commodity acoustic devices. The specifications of employed acoustic devices are listed in Table I. We use both commercial off-the-shelf external speaker–microphone pair and the bottom microphone in

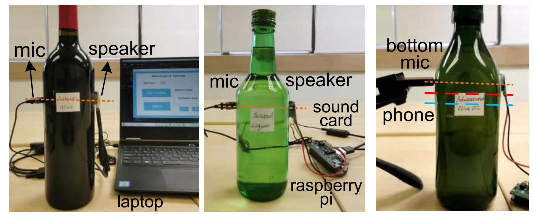


Fig. 17. System setup using external acoustic devices and smartphone.

the smartphone as acoustic devices. As shown in Fig. 17, the acoustic devices are stuck to the two sides of the container surface using the adhesive type. The speaker and microphone are placed horizontally in the middle of the two sides of the container. The acoustic devices can also be flexibly placed at different positions relative to the container, e.g., the red- and blue-dashed lines in the third subfigure in Fig. 17, for more convenient use of our system. External speaker–microphone pair is connected to the raspberry pi/laptop via a common sound card to send and receive the acoustic signal, respectively. The prices of external speakers and microphones are U.S. \$5–\$15. Apart from external acoustic devices, smartphones can also be used. For example, we can use the bottom speaker of the smartphone to send out the acoustic signal, as shown in Fig. 17.

B. Software

We use Python to generate and process the acoustic signal. The generated acoustic signal is made of 21 sine waves whose frequencies range from 18 to 20 kHz with the same interval of 100 Hz. The signal is saved as a WAV file. The sampling rate of acoustic signal is 48 kHz, and the time duration for FFT is 4 s. Then, the frequency resolution after FFT is 0.25 Hz, which is fine-grained enough to extract the amplitude on each desired integer frequency. The VAE and MLP models are trained via PyTorch on a server equipped with Intel Xeon CPU E5-2680 v2 and Nvidia GeForce RTX 2080 GPU with 32-GB memory. When training the model, we use the Adam optimizer and set the learning rate = 1e-4 and betas = (0.9, 0.999). The trained models are stored in the server for detecting the unknown liquid. To evaluate the performance of liquid fraud detection and liquid adulteration ratio recognition, the following metrics are used:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}$$
$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}, \quad \text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$
$$\text{F1 score} = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}. \tag{3}$$

C. Liquid Data Collection

We collect data from liquids of common liquid fraud cases in people's daily life, including liquor, extra-virgin olive oil, wine, and honey frauds, as listed in Table II.

Authentic and Tainted Liquor: High-quality liquor is popular in many countries as daily drinks and gifts. Due to its high price, mixing authentic liquor with cheap and inedible

TABLE II
LIST OF AUTHENTIC AND FAKE LIQUIDS FOR DIFFERENT CASES

Case	Liquids
liquor fraud detection	Authentic: Grey Goose Vodka (GGV, 40%)
	Fake: GGV + random e./m./i., Stolichnaya Vodka (40%)
liquor ratio recognition	GGV + (30%,35%,40%,45%) i.
olive oil fraud detection	Authentic: Colavita extra-virgin olive oil (Ceoo)
	Fake: Ceoo + random c./p./s.
olive oil ratio recognition	Ceoo + (25%,35%,70%,80%) c.
wine fraud detection (different brands)	Authentic: Torres Mas La Plana
	Fake: Barefoot, Penfolds, Mirassou
wine fraud detection (different grapes)	Authentic: Barefoot (Pinor Noir)
	Fake: Barefoot (Shiraz, Merlot, Zinfandel)
honey fraud detection	Authentic: Comita Manuka honey
	Fake: wildflower, lemon, longan flower honey
honey MGO recognition	Comita Manuka honey (MGO levels: 83+, 263+, 514+)

alcohol for sale is the main method of liquor counterfeiting. Therefore, we prepare a high-quality authentic liquor product, the Grey Goose Vodka, and other kinds of alcohol, including ethanol (e.), methanol (m.), and isopropanol (i.), to mix with the authentic liquor as fake liquids. To show the system's ability to detect fake liquids with random mixtures of fraudulent liquids, we mix the authentic liquor with ethanol, methanol, and isopropanol, respectively. For each fraudulent alcohol, we make three bottles of fake liquids, which are obtained by randomly mixing the liquor with the corresponding alcohol. To evaluate the system for detecting the fake liquid, which has the same alcohol level as the authentic liquor but with lower quality, we choose a cheaper vodka, Stolichnaya Vodka, as the fake liquor. We also mix the authentic liquor with different percentages of the isopropanol, including 30%, 35%, 40%, and 45%, to evaluate the performance of adulteration ratio recognition.

Authentic and Adulterated Extra-Virgin Olive Oil: Extra-virgin olive oil provides many nutrients and antioxidants that are beneficial to people's health. While the price of extra-virgin olive oil is usually eight to ten times higher than the prices of canola oil or soybean oil. Some oil sellers would mix the real extra-virgin olive oil with cheaper oil to make more profits. Thus, we prepare an authentic olive oil, the Colavita extra-virgin olive oil, and other kinds of cheap oil, including canola (c.) oil, soybean (s.) oil, and peanut (p.) oil. We first randomly mix the authentic oil with the canola oil, soybean oil, and peanut oil, respectively. Meanwhile, we mix the authentic olive oil with different percentages of the canola oil, including 50%, 60%, 70%, and 80%.

Wine Fraud: Relabeling cheap wines to expensive ones is a common wine fraud. The expensive wines can be simply counterfeited by changing the wine label. In this case, we prepare an expensive wine, Torres Mas La Plana (grape: Cabernet Sauvignon), as the authentic wine, and three cheap wines, Barefoot California (grape: Merlot), Penfolds Koonunga Hill (grape: Shiraz), and Mirassou California (grape: Pinot Noir), with different grape types and brands as the fake liquids. In addition, to test whether the wine with different grape types can be detected, we prepare four wines of the same brand Barefoot California but different grape types, including Pinot Noir, Shiraz, Merlot, and Zinfandel. Pinot Noir is regarded as the authentic wine and the other three types of grapes are treated as fake wines against Pinot Noir.

Honey Fraud: Honey is the third most faked food in the world. The quality of honey varies a lot. The honey with a higher level of methylglyoxal (MGO) is much more expensive than ordinary honey. It is common that high-quality honey is replaced with poor one for sales. Thus, we prepare one high-quality honey, Comvita Manuka Honey with MGO, as the authentic honey, and select three cheaper honey, i.e., wildflower honey, lemon honey, and longan flower honey, as fake honey. In addition, to show the system's ability to recognize the Manuka honey with different MGO levels, we prepare three MGO levels Manuka honey (83+, 263+, and 514+).

During data collection, we measure the acoustic signal at different device-container positions to train and test the model, where ten traces of the acoustic signal are collected at each position. For each case of liquid fraud detection, we use 70 manually collected AATCs from the authentic liquid to train the data augmentation model and generate 400 more AATCs, which are combined with the 70 manually collected AATCs to train the fraud detection model. Then, we collect another 100 AATCs from each of the authentic and fake liquids at random positions to test the model. For adulteration ratio classification, 50 AATCs are manually collected from each ratio of liquids, which are separately augmented with 400 more AATCs. The classification model is trained with the manually collected and VAE generated AATCs of all the liquids. Finally, we collect another 100 AATCs for each ratio of liquids at random positions to test the classification model.

D. Evaluation Results

1) Overall Performance: First, we show the overall performance of the system on liquid fraud detection and adulteration ratio recognition. In Fig. 18, the accuracy, precision, recall, and F1 score are shown for all the liquid fraud cases. The accuracy of liquid fraud detection is around 92%–96%. Specifically, for the liquor (liquor+$\mathbf{x}$(r), $\mathbf{x}$: ethanol, methanol, or isopropanol) and olive oil (olive+$\mathbf{x}$(r), $\mathbf{x}$: canola, soybean, or peanut oil) fraud detection, the average accuracy is approximately 95%, showing that the system can accurately detect the fake liquids with random mixtures of fraudulent liquids. Meanwhile, the system can detect the fake cheap liquor whose alcohol level is the same as the expensive authentic liquor with an accuracy of 92%. For wine fraud detection, the accuracy when detecting the fake wines whose brands and grape types are all different from the authentic wine is about 96%. While the accuracy drops a little to around 93% for detecting the fake wines whose brands are the same but with different grape types. The accuracy of detecting honey fraud is about 95%. We also prepare two bottles of honey with a similar MGO level (263+) but different brands to investigate whether our method

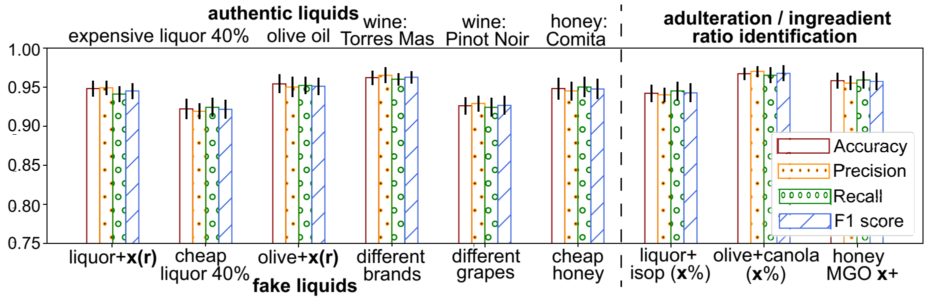


Fig. 18. Overall performance for liquid fraud detection and adulteration ratio recognition for different liquid fraud cases.

can differentiate similar kind of liquids but produced by different companies. We measure the AATCs from one bottle of honey (regard as the authentic honey) and train the liquid fraud detection model and use the AATCs collected from the other bottle of honey (regard as the fake honey) to test the model. The results show that 89.3% of the testing samples are accurately detected as the fake honey. This is because, although having the same MGO level, they are still different in other ingredients, e.g., the amount of carbohydrate and sugar is different. This indicates that the components of the two types of honey still have some differences so that their absorption of the acoustic signal would be distinctive.[3]

For liquid adulteration ratio recognition, we train the ratio classification models for the isopropanol-mixed liquor and canola-mixed olive oil with four different adulteration ratios, respectively. As shown in the right part of Fig. 18, the accuracy for recognizing the adulteration ratio of liquor with the ratio difference of 5% is around 94%, and the recognition accuracy of the olive oil adulteration ratio with 10% interval can reach about 97%. Besides, the Manuka honey with different levels of MGO can be recognized with an accuracy of 95%.

2) Impact of AATC Granularity: In Section III-B, we mentioned that the frequency interval I_f determines the granularity of AATC. A smaller I_f can result in a more fine-grained AATC, which has more frequencies in the AATC. The AATC granularity may affect the liquid fraud detection accuracy. Therefore, we change I_f ranging from 50 to 300 Hz, which results in 7 ($I_f = 300$ Hz), 11 ($I_f = 200$ Hz), 21 ($I_f = 100$ Hz), and 41 ($I_f = 50$ Hz) frequencies in the range of [18 kHz, 20 kHz]. The average accuracy of using different numbers of frequencies for all the liquid fraud cases is shown in Fig. 19. The results show that when the number of frequencies is less than 21, the detection accuracy increases with the growing number of frequencies. This is because more number of frequencies involves more information of the acoustic absorption and transmission in the liquid. However, the accuracy slightly drops for 41 frequencies. Meanwhile, we compare the $F1$ score using different I_f, and $I_f = 100$ Hz with 21 frequencies also achieves the highest $F1$ score of 94.6%. This may due to the reason that more frequencies introduce more redundant and noisy information in the AATC, which could lead to misdetection of the liquid. Therefore, in our system, we use 21 frequencies with an interval of 100 Hz to generate the acoustic signal.

[3]We find our method does not work when the honey experiences crystallization because the honey components change in this process.

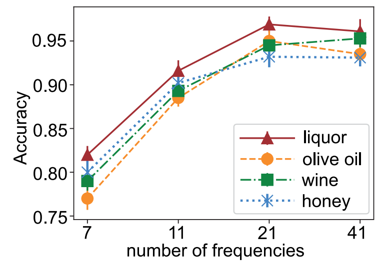


Fig. 19. Accuracy of liquid fraud detection with different numbers of frequencies in AATC.

3) Performance With Different Acoustic Devices: In this evaluation, we show the performance of the system for liquid fraud detection using different acoustic devices. First, we use two different sets of acoustic devices, including external speaker and microphone, and external speaker and bottom microphone of the smartphone, to train and test the liquid fraud detection model, respectively. When using both external speaker and microphone, the accuracy and $F1$ score are 0.956 and 0.954, respectively. When using the external speaker and the smartphone's bottom microphone, the accuracy and $F1$ score are 0.935 and 0.934, respectively. Second, we use different external speakers and microphones to evaluate our calibration method. We use three speakers (S1, S2, and S3) and two microphones (M1 and M2) to evaluate our proposed AATC calibration method for removing the effect of hardware diversity. S1 and S2 are from the same brand (B1) but with different specifications, while S3 is from another brand (B2). M1 and M2 are from different brands B3 and B4, respectively. In the experiment, we first use speaker B1-S1 and microphone B3-M1 to collect both the training and the first testing data. Then, to evaluate the detection accuracy of using different devices from the same brand, we keep the microphone B3-M1 while using another speaker B1-S2 to collect the second testing data. Furthermore, we use speaker B2-S2 and microphone B4-M2 to collect the third testing data to evaluate the performance of liquid fraud detection using different brands' devices. The results are shown in Fig. 20. Comparing with the accuracy of using the training and testing data from the same acoustic devices, the accuracy of using different

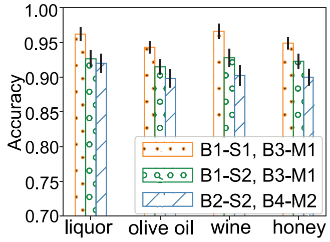


Fig. 20. Accuracy of liquid fraud detection (different acoustic devices).

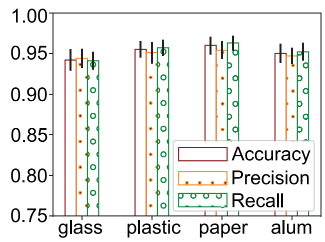


Fig. 21. Liquid fraud detection using different containers.

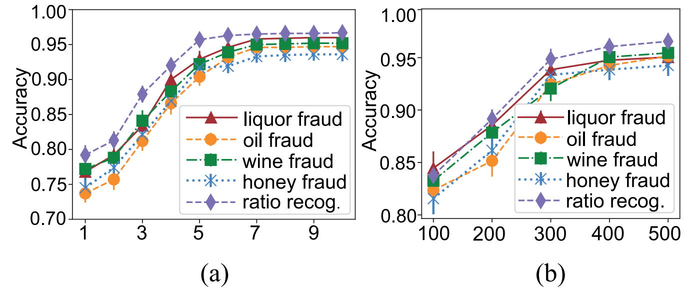


Fig. 22. Effect of the number of positions to collect the AATCs for training the VAE model and the number of augmented AATCs for liquid detection. (a) Number of positions. (b) Number of generated AATCs.

acoustic devices with the same brand still exceeds 92% with only 3%–4% decrease of accuracy, which shows the effectiveness of our AATC calibration method. The accuracy further decreases slightly to about 89%–91% when using devices from different brands because the frequency responses of different brand's devices have larger deviations of the measured AATCs.

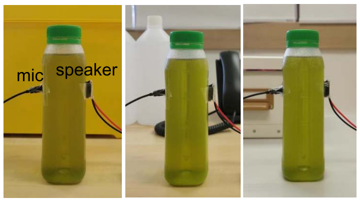
4) Impact of Liquid Container Material: To show the effectiveness of our system on liquid fraud detection using different container materials, we use glass, plastic, paper, and aluminum containers to fill in the authentic and fake liquor, respectively. For each kind of container, we train a separate anomaly detection model using the AATCs of the authentic liquor and then test the model using the AATCs of the authentic and fake liquor. The average accuracy, precision, and recall for different containers are shown in Fig. 21, which all exceed 90%. This article container has slightly higher accuracy compared with that of the glass container because this article container is thinner than the glass container, which incurs less impact on AATC extraction. Besides, our system achieves around 95% accuracy using the aluminum container, which cannot be used by RF-based methods. This is because the RF signal transmission could be significantly affected by the metal. In practice, metal can reflect most of the RF signal, and the RF signal attenuates very fast in the liquid, which results in an extremely weak received signal. In addition, metal materials could change the hardware property of the RF devices (e.g., RFID tag's impedance), making the RF signal undetectable.

5) Impact of AATC Augmentation: When using VAE, two factors can affect the liquid detection performance. The first factor is the number of relative device-container positions to collect the AATCs for training the VAE model. The second factor is the number of generated AATCs from VAE for training the anomaly detection and ratio recognition models. In this evaluation, we test the system performance for each factor.

First, we investigate the effect of different numbers of positions to collect the AATCs for training the VAE model. We choose ten different positions and use different numbers of them (from 1 to 10) to collect AATCs and train the VAE model. Then, we ask volunteers to randomly collect AATCs from another ten positions to test the model. We guarantee that the positions for training and testing do not overlap. At each position, ten AATCs are collected. The AATCs generated from the VAE model are mixed with the manually collected AATCs to train the liquid detection models. We fix the number of generated AATCs from VAE to 400. The accuracy of using the AATCs from different numbers of positions is shown in Fig. 22(a). With more positions' AATCs to train the VAE model, the testing accuracy gradually increases. This is because the VAE model can learn more patterns from more device-container positions. In our experiment, for liquid fraud detection, the accuracy does not improve obviously after the number of trained positions exceeds 7. Therefore, we only need to collect the training AATCs from seven positions (i.e., 70 AATCs), and the AATCs collected from other positions can be accurately detected. For ratio recognition, the AATCs from five different positions (i.e., 50 AATCs) are collected from each ratio of liquid to train the classification model, which already achieves an average classification accuracy of around 96%.

Second, we investigate the effect of different numbers of generated AATCs from VAE. We fixed the number of manually collected AATCs to 70 (for anomaly detection) and 50 (for ratio recognition) while using 100, 200, 300, 400, and 500 generated AATCs, which are combined with manually collected AATCs, to train liquid detection models. As shown in Fig. 22(b), the accuracy increases with more number of generated AATCs. When the number of generated AATCs reaches 300, the average accuracy for all cases exceeds 90%. As the number increases to 400, the average accuracy exceeds 94%. When the number is above 400, our system shows no pronounced improvement. Thus, we generate 400 AATCs from the VAE to augment the training data.

Third, we compare the liquid detection performance without VAE, with VAE using original loss function, with VAE using our new loss function. First, we train the anomaly detection model and adulteration ratio classification model only using manually measured AATCs without VAE for data augmentation. Then, we train two liquid detection models with manually measured and augmented AATCs using the original VAE loss function. Finally, we use VAE and the new loss function to generate AATCs and train the models. All models



(a) (b) (c)

Fig. 23. Different tables and surrounding layouts. (a) Layout 1. (b) Layout 2. (c) Layout 3.

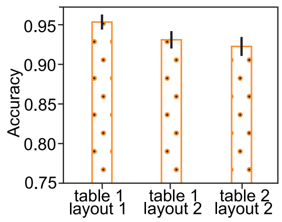
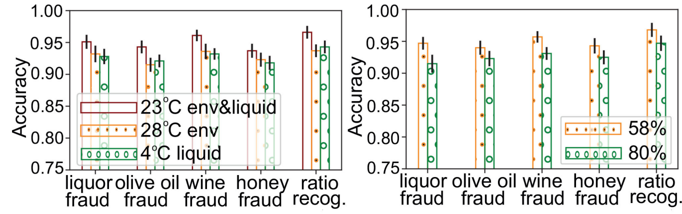


Fig. 24. Detection results under different tables and layouts.

are tested using the AATCs randomly collected from different device-container positions. The detection results show that the accuracy without augmentation is around 10% lower than that using the VAE and original loss function. By applying our new loss function, the accuracy can be further improved by another 3%–5%. Therefore, our dedicatedly designed VAE model can indeed augment effective AATCs for different positions to train the liquid detection models.

6) Impact of different Environments: In practice, the acoustic signal could be reflected by surrounding objects, which would be received by the microphone. As a result, different tables and surrounding objects could bring multipath signals in the received acoustic signal. In this evaluation, we change the table and surrounding objects around the container to investigate whether the detection performance would be affected by different tables and layouts. We first place liquids on Table I (made of wood) with layout 1 as shown in Fig. 23(a) to measure the AATC and train and test the liquid detection model. Then, we apply the trained model to detect the same liquids but under different tables (Table II: made of metal) and layouts (layout 2 and layout 3: different surrounding objects), as shown in Fig. 23(b) and (c). The detection results are given in Fig. 24. The accuracy all exceeds 90% and only decreases 2%–4% under different tables and layouts. This is because we intentionally place the microphone in the acoustic shadow zone, within which the table reflected signals and other diffracted signals are significantly reduced [21]. Meanwhile, we design a data augmentation method using VAE to help improve the robustness of our system in the face of multipath signals.

The temperature and humidity can influence the acoustic impedance, which may affect the system performance. Thus, we change the temperature and humidity to see their effects on liquid fraud detection and ratio recognition. To investigate the temperature effects, we first collect AATCs with a 23 °C temperature for both the environment and liquid to train and test liquid detection models. Then, we keep the temperature



(a) (b)

Fig. 25. Accuracy of liquid fraud detection and adulteration ratio recognition with different temperature and humidity ratios. (a) Accuracy for different temperature. (b) Accuracy for different humidity.

of the liquid while increasing the environment temperature to 28 °C to test the models. Next, we keep the environmental temperature while using the same liquid with a 4 °C temperature to test the models. For the humidity effect, we first collect the training and testing AATCs in a 58% humidity ratio environment. Then, we increase the humidity ratio to around 80% and collect a set of testing AATCs. The accuracy of liquid fraud detection and ratio recognition under different temperatures and humidity ratios is shown in Fig. 25. The results show that the accuracy under the 28 °C environment temperature, 4°C liquid temperature, and the 80% humidity ratio is comparable with those under the 23 °C environment and liquid temperature as well as 58% humidity ratio, respectively. This is because temperature and humidity impact the velocity v of the acoustic signal in a linear way [22], [31], and v is generally inversely proportional to the acoustic absorption ratio of liquid α in the frequency band ranging from 18 to 20 kHz, i.e., $\alpha \sim 1/v$ [32]. Thus, AATC changes linearly with different temperatures and humidity ratios within the same frequency band, which can be removed by the AATC normalization step.

V. Related Work

In this section, we introduce the related works for liquid detection using different sensors and signals.

A. Chemical and Biological Sensor-Based Systems

Chemical and biological sensors have been used to identify the target analytes in the liquid by food labs and industries. They aim to extract properties of the biomolecules via various methods, e.g., electrochemical [5] and mass-based detection [33]. In food and liquid testing labs, many tools, e.g., the infrared spectrometer, are used to detect and analyze various kinds of analytes and contaminants in the liquid. The chemical and biological sensor-based systems are generally expensive and require complicated operations for measuring liquid properties, which are unavailable for public use. In addition, the chemical and biological sensors require direct contact with the liquid, which is quite intrusive and inconvenient for sealed liquids. In our work, we propose a liquid fraud detection system using cost-effective acoustic devices, which can detect the fake liquids in a nonintrusive manner without opening the liquid container.

B. RF Signal-Based Systems

Recently, researchers use RF signals to detect liquids [9]–[13]. They employ the properties of the dielectric spectroscopy or the impedance of the liquid to differentiate different liquids. RF-EAT [9] and RFIQ [10] enable liquid and food sensing by measuring the near-field coupling effect between RFID tags and the liquid. LiquID [11] uses an UWB radar signal to estimate the liquid permittivity by measuring the ToF of the radar signal traveling through the liquid. However, these systems require specialized devices (e.g., USRP and UWB radar), which is difficult for public use. Tagtag [12] and TagScan [13] leverage commercial off-the-shelf RFID devices to measure the changes of the RFID signal's phase and RSSI based on the impedance change of the tag caused by different liquids. However, RF signals cannot detect the liquids with metal containers. For the RFID-based method, metal containers could significantly affect the RFID tag's impedance [34], which makes the RFID tag undetectable. Hence, it cannot sense the liquid inside the container. UWB radar leverages the RF signal traveling through the liquid for liquid detection. However, metal containers reflect most of the RF signal, and the RF signal attenuates significantly in the liquid, resulting in an extremely weak received signal. In contrast, our acoustic-based method can work properly for metal containers, so that our system can be applied to cope with more kinds of containers.

C. Acoustic-Based Systems

Ultrasound sensors are employed for detecting the contaminants in the liquid by measuring the tiny penetration depth of the shear waves [35]. Ultrasound sensors can measure the sound speed in liquid for liquid detection since the acoustic signal travels at different speeds in different liquids. However, the ultrasound-based system requires wide frequency bandwidth (e.g., >500 kHz) to accurately measure the ToF, which is not supported by most commodity acoustic devices. Compared with ultrasound-based systems, our system only uses commodity acoustic devices to measure the AATC, which is more cost effective. The acoustic reflection is also employed for liquid-related applications. SoQr [36] employs the liquids reflection of the acoustic signal and extracts the Mel-frequency cepstral coefficients to train a liquid level classification model. However, such a system needs to open the liquid container to expose the liquid surface for measuring the acoustic reflection. Different from SoQr, our system investigates the liquids intrinsic characteristic, i.e., acoustic impedance, which is revealed in the absorption of acoustic signal. Such an important feature enables differentiating fake liquids from the authentic one. Meanwhile, our method does not require opening the liquid container, which provides a nonintrusive way for liquid detection.

D. Other Systems

QET-based systems detect flammable and explosive liquids, which have been widely deployed at many public places, such as the airport and train station [6]. QET technique measures electrical properties of the liquid, i.e., permittivity and conductivity, to test whether the liquid is flammable. While current QET-based systems cannot realize liquid fraud detection. Another kind of system uses the tensiometer [7] or the camera [8] to measure the surface tension of the liquid to identify the liquid type. However, they need to open the liquid container for measuring the tension, and the tensiometer could cost thousands of U.S. dollars.

VI. Discussion

We discuss several practical issues about using the HearLiquid system in this section. First, there could be many background noises, which may mix with the measured acoustic signal. However, frequencies of most background noises in the environments, as well as the human voice, are lower than 8 kHz. In our work, to make the sound inaudible, we select the frequency band of [18 kHz, 20 kHz] for the acoustic signal. We note that the gap between the frequency band we apply and the background noise is as far as 10 kHz. Therefore, in our work, we remove background noises in the environment with a high-pass filter.

Second, in HearLiquid, the extracted AATC of the liquid, in fact, involves the effect of the liquid container. While, the containers of liquids are usually the same for the same category of liquids since counterfeiting and adulterated liquids usually use the same container as the authentic liquid to cheat consumers. Thus, the container effect can be ignored in the current system. However, when the liquid container changes, the detection model needs to be updated with the AATCs collected with the new container. Although it is not complicated to upgrade the model, it would be better if we can remove the effect of container. In future work, we will try to deal with the containers effect with the transfer learning technique, which may transfer the liquid's property measured with one container to another container with less data collection and model training effort.

VII. Conclusion

In this work, we proposed HearLiquid, which can detect the liquid fraud in a nonintrusive manner using commodity acoustic devices. It is the first time that commodity acoustic devices are used for liquid detection. We extract the AATC from the acoustic signal traveling through the liquid, i.e., AATC, which can be used to differentiate different liquids. To deal with the practical factors for extracting the AATC, we proposed a series of methods to tackle the hardware diversity of the acoustic devices and the AATC variations brought by different relative device-container positions. We applied a reference signal to calibrate different frequency responses caused by the hardware diversity. Based on the patterns of the measured AATCs at different relative device-container positions, we leveraged a data augmentation technique to automatically emulate a large number of AATCs under different positions. The augmented AATCs sufficiently copes with the AATC variations to promote the liquid detection accuracy. We conduct extensive experiments on various liquid fraud cases under different experiment settings. The experimental results show that HearLiquid achieves an overall accuracy up to 97%.

References

[1] "Counterfeit Foods, Illegally Labelled and Grey Market Goods: Is Your Brand Protected?" Global Food Safety Resource. 2020. [Online]. Available: https://globalfoodsafetyresource.com/grey-markets-products/ (accessed May 1, 2020).

[2] N. N. Misra, Y. Dixit, A. Al-Mallahi, M. S. Bhullar, R. Upadhyay, and A. Martynenko, "IoT, big data and artificial intelligence in agriculture and food industry," *IEEE Internet Things J.*, early access, May 29, 2020, doi: 10.1109/JIOT.2020.2998584.

[3] S. Mondal, K. P. Wijewardena, S. Karuppuswami, N. Kriti, D. Kumar, and P. Chahal, "Blockchain inspired RFID-based information architecture for food supply chain," *IEEE Internet Things J.*, vol. 6, no. 3, pp. 5803–5813, Jun. 2019.

[4] S. N. Jha, P. Jaiswal, M. K. Grewal, M. Gupta, and R. Bhardwaj, "Detection of adulterants and contaminants in liquid foods—A review," *Crit. Rev. Food Sci. Nutrition*, vol. 56, no. 10, pp. 1662–1684, 2016.

[5] S. Viswanathan, H. Radecka, and J. Radecki, "Electrochemical biosensors for food analysis," *Monatshefte fur Chemie Chemical Monthly*, vol. 140, no. 8, p. 891, 2009.

[6] Q. Marashdeh, L.-S. Fan, B. Du, and W. Warsito, "Electrical capacitance tomography—A perspective," *Ind. Eng. Chem. Res.*, vol. 47, no. 10, pp. 3708–3719, 2008.

[7] F. Tamm, G. Sauer, M. Scampicchio, and S. Drusch, "Pendant drop tensiometry for the evaluation of the foaming properties of milk-derived proteins," *Food Hydrocolloids*, vol. 27, no. 2, pp. 371–377, 2012.

[8] S. Yue and D. Katabi, "Liquid testing with your smartphone," in *Proc. ACM MobiSys*, 2019, pp. 275–286.

[9] U. Ha, J. Leng, A. Khaddaj, and F. Adib, "Food and liquid sensing in practical environments using RFIDs," in *Proc. USENIX NSDI*, 2020, pp. 1083–1100.

[10] U. Ha, Y. Ma, Z. Zhong, T.-M. Hsu, and F. Adib, "Learning food quality and safety from wireless stickers," in *Proc. ACM HotNets*, 2018, pp. 106–112.

[11] A. Dhekne, M. Gowda, Y. Zhao, H. Hassanieh, and R. R. Choudhury, "LiquID: A wireless liquid identifier," in *Proc. ACM Mobisys*, 2018, pp. 442–454.

[12] B. Xie *et al.*, "TagTag: Material sensing with commodity RFID," in *Proc. ACM SenSys*, 2019, pp. 338–350.

[13] J. Wang, J. Xiong, X. Chen, H. Jiang, R. K. Balan, and D. Fang, "TagScan: Simultaneous target imaging and material identification with commodity RFID devices," in *Proc. ACM MobiCom*, 2017, pp. 288–300.

[14] J. J. Markham, R. T. Beyer, and R. B. Lindsay, "Absorption of sound in fluids," *Rev. Mod. Phys.*, vol. 23, no. 4, p. 353, 1951.

[15] C. M. Davis Jr. and J. Jarzynski, "Liquid water—Acoustic properties: Absorption and relaxation," in *The Physics and Physical Chemistry of Water*. Boston, MA, USA: Springer, 1972, pp. 443–461.

[16] P. M. Morse, R. H. Bolt, and R. L. Brown, "Acoustic impedance and sound absorption," *J. Acoust. Soc. Amer.*, vol. 12, no. 2, pp. 217–227, 1940.

[17] J.-P. Dalmont, "Acoustic impedance measurement, part I: A review," *J. Sound Vib.*, vol. 243, no. 3, pp. 427–439, 2001.

[18] H. Lee, T. H. Kim, J. W. Choi, and S. Choi, "Chirp signal-based aerial acoustic communication for smart devices," in *Proc. IEEE INFOCOM*, Hong Kong, 2015, pp. 2407–2415.

[19] A. London, "The determination of reverberant sound absorption coefficients from acoustic impedance measurements," *J. Acoust. Soc. Amer.*, vol. 22, no. 2, pp. 263–269, 1950.

[20] B. Taylor and H. G. Mueller, *Fitting and Dispensing Hearing Aids*. San Diego, CA, USA: Plural Publ., 2016.

[21] J. Piechowicz, "Sound wave diffraction at the edge of a sound barrier," *Acta Physica Polonica A*, vol. 119, no. 6A, pp. 1040–1045, 2011.

[22] P. A. Oliveira, R. D. M. B. Silva, G. C. D. Morais, A. V. Alvarenga, and R. P. B. Costa-Félix, "Speed of sound as a function of temperature for ultrasonic propagation in soybean oil," *J. Phys. Conf. Ser.*, vol. 733, no. 1, 2016, Art. no. 012040.

[23] S. Rosen and P. Howell, *Signals and Systems for Speech and Hearing*, vol. 29. Leiden, The Netherlands: Brill, 2011.

[24] C. E. Y. Dorfan, S. Gannot, and P. A. Naylor, "Speaker localization with moving microphone arrays," in *Proc. EUSIPCO*, 2016, pp. 1003–1007.

[25] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge, U.K.: Cambridge Univ. Press, 2005.

[26] A. A. Cook, G. Mısırlı, and Z. Fan, "Anomaly detection for IoT time-series data: A survey," *IEEE Internet Things J.*, vol. 7, no. 7, pp. 6481–6494, Jul. 2020.

[27] Y. Wang, J. Shen, and Y. Zheng, "Push the limit of acoustic gesture recognition," in *Proc. IEEE INFOCOM*, Toronto, ON, Canada, 2020, pp. 566–575.

[28] A. A. Pol, V. Berger, C. Germain, G. Cerminara, and M. Pierini, "Anomaly detection with conditional variational autoencoders," in *Proc. IEEE ICMLA*, Boca Raton, FL, USA, 2019, pp. 1651–1657.

[29] H. Xu *et al.*, "Unsupervised anomaly detection via variational autoencoder for seasonal KPIs in Web applications," in *ACM WWW*, 2018, pp. 187–196.

[30] F. Pukelsheim, "The three sigma rule," *Amer. Stat.*, vol. 48, no. 2, pp. 88–91, 1994.

[31] W. Wilson and D. Bradley, "Speed of sound in four primary alcohols as a function of temperature and pressure," *J. Acoust. Soc. Amer.*, vol. 36, no. 2, pp. 333–337, 1964.

[32] Z. Mohamed and L. Egab, "Tire cavity noise mitigation using acoustic absorbent materials," in *Automotive Tire Noise and Vibrations*. Kidlington, U.K.: Elsevier, 2020, pp. 245–270.

[33] I. Mannelli, M. Minunni, S. Tombelli, and M. Mascini, "Quartz crystal microbalance (QCM) affinity biosensor for genetically modified organisms (GMOs) detection," *Biosens. Bioelectron.*, vol. 18, nos. 2–3, pp. 129–140, 2003.

[34] J. T. Prothro, G. D. Durgin, and J. D. Griffin, "The effects of a metal ground plane on RFID tag antennas," in *Proc. IEEE Antennas Propag. Soc. Int. Symp.*, 2006, pp. 3241–3244.

[35] E. E. Franco, J. C. Adamowski, and F. Buiochi, "Ultrasonic sensor for the presence of oily contaminants in water," *Dyna*, vol. 79, no. 176, pp. 4–9, 2012.

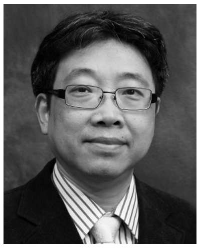
[36] M. Fan and K. N. Truong, "SoQr: Sonically quantifying the content level inside containers," in *Proc. ACM Ubicomp*, 2015, pp. 3–14.

Yanni Yang received the B.E. and M.Sc. degrees from the Ocean University of China, Qingdao, China, in 2014 and 2017, respectively, and the Ph.D. degree in computer science from the Hong Kong Polytechnic University, Hong Kong, in 2021.

She is currently a Postdoctoral Fellow with the Department of Computing, The Hong Kong Polytechnic University. She visited the Media Lab in Massachusetts Institute of Technology, Cambridge, MA, USA, in 2019, as a visiting student. She has published papers in many conferences and journals, such as ACM Ubicomp, IEEE SECON, IEEE/ACM TRANSACTIONS ON NETWORKING, IEEE TRANSACTIONS ON MOBILE COMPUTING, and IEEE INTERNET OF THINGS. Her research interests include wireless human sensing, pervasive and mobile computing, and Internet of Things.

Yanwen Wang (Member, IEEE) received the B.S. degree in electronic engineering from Hunan University, Changsha, China, in 2010, the M.S. degree in electrical engineering from the Missouri University of Science and Technology, Rolla, MO, USA, in 2013, and the Ph.D. degree from the Department of Computing, Hong Kong Polytechnic University, Hong Kong, in 2020.

He is currently an Associate Professor with the College of Electrical and Information Engineering, Hunan University. His research interest includes mobile and network computing, RFID systems, and acoustic sensing.



Jiannong Cao (Fellow, IEEE) received the M.Sc. and Ph.D. degrees in computer science from Washington State University, Pullman, WA, USA, in 1986 and 1990, respectively.

He is currently the Chair Professor with the Department of Computing, Hong Kong Polytechnic University (PolyU), Hong Kong, where he is also the Dean of Graduate School and the Director of Research Institute of Artificial Intelligent of Things, the Internet and Mobile Computing Lab, and the Vice Director of the University's Research Facility in Big Data Analytics. He has co-authored five books, co-edited nine books, and published over 500 papers in major international journals and conference proceedings. His research interests include distributed systems and blockchain, wireless sensing and networking, big data and machine learning, and mobile cloud and edge computing.

Dr. Cao is a member of Academia Europaea and an ACM distinguished member.

Jinlin Chen received the master's degree from the Hong Kong Polytechnic University, Hong Kong, in 2016, where he is currently pursuing the Ph.D. degree with the Department of Computing.

From 2017 to 2018, he works as a Software Engineer with the Hong Kong Applied Science and Technology Research Institute. He has published papers in many conferences and journals, such as *Knowledge-Based Systems*, IEEE ICPADS, and IEEE ICCCN. His research interests include distributed multirobot systems, robotic middleware, and cooperative robotic system control.