LiquImager: Fine-grained Liquid Identification and Container Imaging System with COTS WiFi Devices

FEI SHANG, University of Science and Technology of China, China

PANLONG YANG*, Nanjing University of Information Science and Technology, China

DAWEI YAN, University of Science and Technology of China, China

SIJIA ZHANG, University of Science and Technology of China, China

XIANG-YANG LI, University of Science and Technology of China, China and Deqing Alpha Innovation Institute, China

WiFi has gradually developed into one of the main candidate technologies for ubiquitous sensing. Based on commercial off-the-shelf (COTS) WiFi devices, this paper proposes *LiquImager*, which can simultaneously identify liquid and image container regardless of container shape and position. Since the container size is close to the wavelength, diffraction makes the effect of the liquid on the signal difficult to approximate with a simple geometric model (such as ray tracking). Based on Maxwell's equations, we construct an electric field scattering sensing model. Using few measurements provided by COTS WiFi devices, we solve the scattering model to obtain the medium distribution of the sensing domain, which is used for identifing and imaging liquids. To suppress the signal noise, we propose *LiqU-Net* for image enhancement. For the centimeter-scale container that is randomly placed in an area of 25 cm × 25 cm, *LiquImager* can identify the liquid more than 90% accuracy. In terms of container imaging, *LiquImager* can accurately find the edge of the container for 4 types of containers with a volume less than 500 ml.

CCS Concepts: • **Human-centered computing** → **Ubiquitous and mobile computing design and evaluation methods**.

Additional Key Words and Phrases: Wireless sensing; Liquid identification; Electromagnetic waves; Complex permittivity.

1 INTRODUCTION

Due to the effectiveness of wireless signals in low-light conditions, wireless-based shape identification finds diverse applications in both industrial production and everyday life [25, 39]. For example, in beverage production and warehouse management, when dealing with the same type of liquid packaged in different containers (e.g.,

---

*Corresponding Author

Authors' addresses: Fei Shang, shf_1998@outlook.com, University of Science and Technology of China, No. 96 Jinzhai Road, Hefei, Anhui, China, 230027; Panlong Yang, plyang@ustc.edu.cn, Nanjing University of Information Science and Technology, No. 219 Ningliu Road, Nanjing, Jiangsu, China, 210044; Dawei Yan, yandw@mail.ustc.edu.cn, University of Science and Technology of China, No. 96 Jinzhai Road, Hefei, Anhui, China, 230027; Sijia Zhang, sxzsj@mail.ustc.edu.cn, University of Science and Technology of China, No. 96 Jinzhai Road, Hefei, Anhui, China, 230027; Xiang-Yang Li, xiangyangli@ustc.edu.cn, University of Science and Technology of China, No. 96 Jinzhai Road, Hefei, Anhui, China, 230027 and Deqing Alpha Innovation Institute, Deqing Geographic Information Town, Huzhou, Zhejiang, China, 313299.
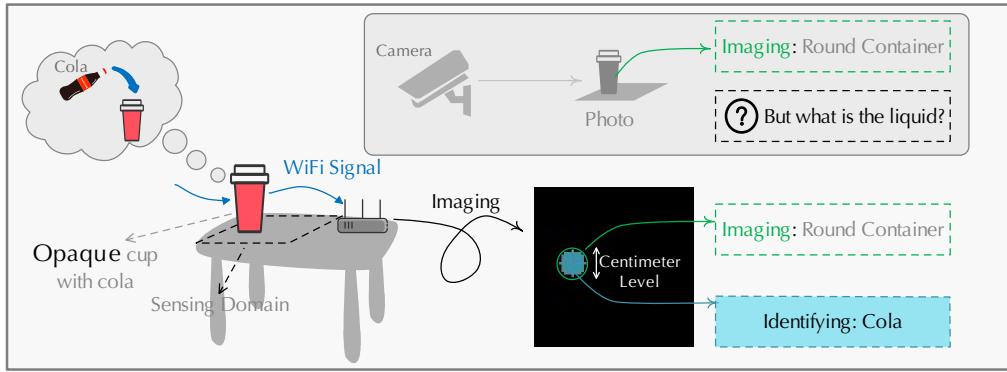
Fig. 1. Container-shape-independent and position-independent liquid identification and centimeter-level containers imaging using commercially available Wi-Fi devices.

large bottles and small bottles), it is essential to apply distinct procedures. With integrated shape identification capabilities into the system, these processes can be carried out more intelligently and efficiently, ensuring appropriate handling and management. Besides, in the context of everyday life, if we can achieve material identification while simultaneously enabling shape identification, it would introduce more exciting features for VR/AR, such as "perspective". This would open up additional possibilities for human-computer interaction.

In recent years, many outstanding works, including LiqRay [33], FGLiquid [25], LiquID [5], and *etc*, have attempted to use wireless signals to identify the material composition of liquids, which offers the potential for ubiquitous liquid sensing. But these excellent works focus more on achieving high-precision identification of liquids in specific containers. However, due to the small size of these containers, diffraction phenomena *make the container shape have a great impact on liquid identification*. In actual scenarios, containers come in various shapes. In addition, its size is close to the wavelength of sub-6G signals. Since slight changes in shape may cause large differences in the received signal, *we need to pay attention to the impact of material and shape on the received signal at the same time*.

In this paper, we propose *LiquImager*, a system that can simultaneously identify liquids and image containers using COTS Wi-Fi devices. Specifically, we have two goals:

- *Imaging of centimeter-level containers.* In the real world, the size of many liquid storage and transportation containers is usually less than 10 cm, such as cups, bottles, and cans. Therefore, in order to cope with different scenarios, we need to image centimeter-scale containers of different shapes.
- *Container-shape-independent and position-independent liquid identification.* In real-world scenarios, the shape of the container and its position relative to the RF devices are often not fixed. Therefore, we need to be able to identify the liquid, whether it is placed at different locations or held in different shaped containers.

Unlike traditional approaches that treat liquid identification and container imaging as two independent tasks, we try to present both material and shape information in one image. Recall that since different items have different colors, we can easily tell whether a green triangle or a blue rectangle is placed on the white table. Similarly, different media (such as water, air, *etc*.) have different complex permittivity [8]. If we obtain the distribution of the permittivity in the sensing domain (the characteristics of different materials in different positions), we can easily complete liquid identification and container imaging through simple processing (such as the k-nearest-neighbor classifier and edge detection). The opportunity to obtain the permittivity distribution comes from a basic observation that the scattered Wi-Fi signal changes with the material, shape, and position of the target, *i.e.*

the channel state information (CSI) of the Wi-Fi signal depends on the permittivity distribution in space [11, 18]. Before CSI is used to infer the complex permittivity distribution of the sensing domain for liquid identification and imaging, three challenges need to be addressed.

(1) *How to describe the complex influence of centimeter-level media on Wi-Fi signals?* The Wi-Fi signal is affected by a combination of the location, shape and type of liquid of the container. However, all of them are unknown to us. To make matters worse, since the wavelength of the Wi-Fi signal is close to the size of the container, the diffraction effect[1] of the Wi-Fi signal is obvious when transmitted near the liquid. As a result, it is difficult to describe it with a simple geometric model (such as ray tracing model).

(2) *How to solve the complex model using few measurements?* We wish to estimate the distribution of complex permittivity in the sensing domain, which requires estimating numerous parameters. The side length of the sensing domain (such as the desktop, luggage, *etc.*) usually exceeds 30 cm, and to complete the imaging of centimeter-level containers, the size of the grid is usually less than 1 cm when discretizing the domain. As a result, the parameters we need to estimate often exceed 1000. As the area increases, the number of parameters to be estimated grows at a rate of $O(n^2)$. Since there are few antennas on commercial Wi-Fi devices (for example, there only three antennas on the Intel 5300 network card), we are often only able to obtain dozens or even a few measurements.

(3) *How to fine-grained identify centimeter-level liquids using noisy data?* The centimeter scale size makes the difference in the effect of different media on the signal very small. For example, when the signal frequency of the electromagnetic wave is 5 GHz, the complex permittivity of mineral water and cola are $73.43 + 17.87j$ and $79.74 + 10.82j$, respectively. When the transmission distance is 1 m, the amplitude of the signal transmitted in water is more than 100 times larger than that transmitted in cola [5]. But when the transmission distance is 1 cm, the former is only 1.57 times of the latter. To make matters worse, the CSI data collected by COTS Wi-Fi devices contains a lot of noise [10, 47, 56], making fine-grained liquid identification more difficult.

Our solutions are as follows. Firstly, in order to describe the complex influence of centimeter-scale media on Wi-Fi signals as accurately as possible, we build an **electric field scattering sensing model** directly from Maxwell's equations (Sec 3). Second, to infer the permittivity distribution of the sensing domain from the received signal, we first model the electric field excited by the Wi-Fi antenna and then incorporate the positional information to estimate the distribution of the incident field. Then combined with our established scattering model, the least squares method is used to estimate the distribution of complex permittivity in the scattering domain. Finally, we take advantage of the noise probability distribution properties to suppress its influence on complex permittivity estimation (Sec. 5). In order to perform high-precision container imaging and fine-grained material identification, we propose a *LiqU-Net* network for image enhancement (Sec. 6).

Contributions: The main contributions in *LiquImager* summarize as follows:

- We design *LiquImager*, which can use COTS Wi-Fi devices to image centimeter-level containers and identify liquids regardless of liquid position and container shape. To the best of our knowledge, this is the first system to simultaneously perform material identification and centimeter-scale target imaging using COTS Wi-Fi devices, which will provide more possibilities for ubiquitous sensing.
- Since the vessel size is similar to the wavelength, diffraction makes the effect of liquid on the signal so complex that it is difficult to approximate it with simple geometric models (such as classical ray-tracing models). Instead, we build an electric field scattering sensing model directly based on Maxwell's equations, which can more accurately describe the influence of the dielectric properties, position, and size of the medium on the signal. Based on this model, we utilize the received signal to infer the complex permittivity distribution in the sensing area, which allows us to solve the problem of container imaging and liquid

---

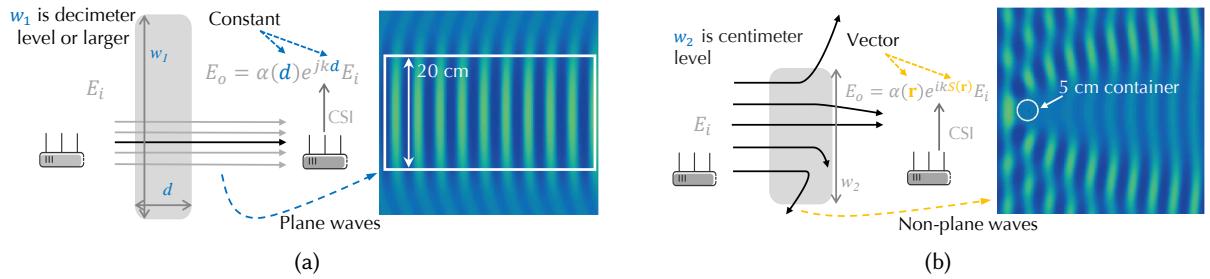[1]The propagation path of the signal is no longer a straight line.

Fig. 2. When the target size is much larger than the wavelength, the influence of the material on the signal is almost completely reflected on one ray (a), while when the target size is close to the wavelength, the material characteristics are difficult to be represented by a simple geometric model (b).

recognition as a whole. Additionally, to eliminate noise and improve the effect of the image, we designed *LiqU-Net* to enhance the image.
- We test *LiquImager* in a area of 25 cm × 25 cm, and the accuracy rate of liquid identification is more than 93.11%. We use 4 different types of containers to hold liquids, and *LiquImager* still has a precision of more than 91% in identifying liquids. In addition, *LiquImager* can accurately image containers arbitrarily placed in the field, and distinguish the shape of the container with an accuracy of 100%.

The rest of the paper is organized as follows. In Section 2, we introduce the complex impact of centimeter-scale containers on Wi-Fi signals. In Section 3, based on Maxwell's equations, we construct a basic electric field scattering sensing model. We introduce the components of our system in Section 4. We detail how to use the received signal for pre-imaging, the method of image enhancement, and how to use the imaging results for container shape detection and liquid identification from Section 5 to Section 7. Implementation and evaluation are presented in Section 8, Section 9 and Section 10. Finally, we discuss the related work in Section 11.

2 WHY RAY-TRACING MODEL DOESN'T WORK WELL FOR CENTIMETER-SCALE MEDIA

Electromagnetic waves (EM waves) traveling in space are described by Maxwell's equations, whose solutions depend on the medium [8]. Traditional material sensing methods [6, 25, 33, 34, 39, 46, 50] often approximate Maxwell's equations as a geometric model (such as rays), and then solve the model by undetermined coefficient to obtain material information. However, *when the size of the target cannot be much larger than the wavelength, it is difficult to seek an accurate geometric approximation.* It is difficult to achieve medium imaging and material identification without accurately describing the influence of the shape, material, and position of the medium on the electric field. Fortunately, we note that the electromagnetic field transmitted in space can be fully described by Maxwell's equations. Therefore, we try to start from Maxwell's equations to explore a new solution.

The propagation characteristics of EM waves vary with materials. Electromagnetic waves propagating in space satisfy Maxwell's equations, which are [8]

$$
\begin{cases}
\nabla \cdot \boldsymbol{E} = \dfrac{1}{\epsilon_0} \rho & \nabla \cdot \boldsymbol{B} = 0 \\
\nabla \times \boldsymbol{E} = -\dfrac{\partial \boldsymbol{B}}{\partial t} & \nabla \times \boldsymbol{B} = \mu_0 \boldsymbol{J} + \mu_0 \epsilon_0 \dfrac{\partial \boldsymbol{E}}{\partial t}
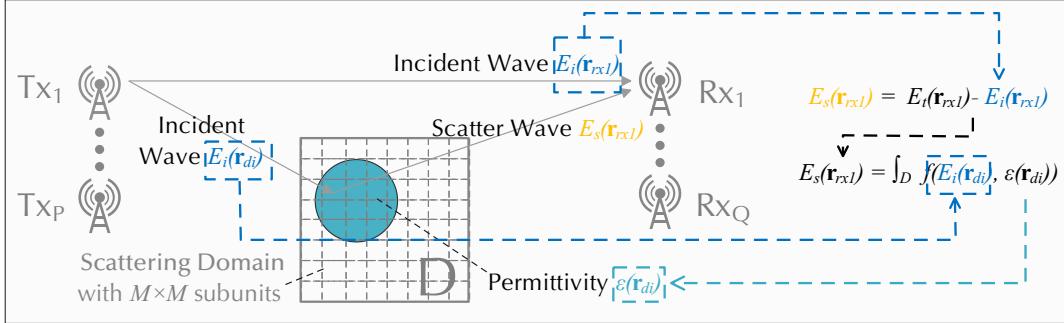\end{cases}
\tag{1}
$$

Fig. 3. Schematic diagram of electric field scattering sensing model. The total electric field $E_t(\mathbf{r}_{rx})$ at the receiving antenna is the superposition of the incident field $E_i(\mathbf{r}_{rx})$ excited by the transmitting antenna and the scattered field $E_s(\mathbf{r}_{rx})$ of the target. The scattered wave $E_s(\mathbf{r}_{rx})$ is a function of the incident wave $E_i(\mathbf{r}_{d_i})$ and the complex permittivity distribution $\epsilon(\mathbf{r}_{d_i})$. We first obtain the scattered waves $E_s(\mathbf{r}_{rx})$ from the measurements, and then solve the distribution of the complex permittivity of the domain D using the least squares method combined with the incident waves $E_i(\mathbf{r}_{d_i})$.

where $\mathbf{E}$ and $\mathbf{B}$ are the electric and magnetic fields, respectively. ϵ_0 and μ_0 are the permittivity and permeability of free space, respectively. ρ is the charge density and $\mathbf{J}$ is the current density, and *both of them are related to the property of molecule and/or atom.* As a result, the solutions of Eq. 1 vary with material.

Traditional wireless sensing methods [6, 25, 33, 34, 39, 46, 50] usually use geometry models to approximate Eq. 1. The most common model is the **ray tracing model**, which equates waves to rays and uses geometric relations to determine the reflection and refraction paths of waves. In particular, as shown in Fig. 2a, for a media with a width is d, the relationship between the outgoing wave E_o and the incoming wave E_i can be expressed as:

$$E_o = \alpha(d)e^{-jkd}E_i, \tag{2}$$

where α is the attenuation coefficient and k is the wavenumber in the media. When the transmission distance d is fixed, the pair (α, k) can fully describe the influence of the material on the wave. For a cylinder whose size is 10 times the wavelength, the electric field around it is shown on the right side of Fig. 2a. In a domain with a width of 20 cm, the wave can be considered as a plane wave. Therefore, for the received signal, we can use a ray to describe the wave transmission.

However, when sensing centimeter-level targets with Wi-Fi signals, it is difficult to approximate them with geometric models. As shown in Fig. 2, when the target is small, the signals are scattered in all directions. For a cylinder whose size is similar to the wavelength, the electric field around it is shown on the right side of Fig. 2b. Scattered waves are very distinctly non-plane waves. As a result, the approximate results of the ray tracing model have large errors at this time. For the Wi-Fi signals whose wavelength is grater than 5 cm, which is close to the container size, the ray tracing model is not suitable.

Since the type, shape, and location of the liquid will affect the received signal, in order to image and identify centimeter-level liquids, *we hope to describe the scattered signal of the liquid as accurately as possible.* This will provide the basis for us to use CSI data to complete perception tasks (from Sec. 5 to Sec. 7). Therefore, we try to establish a more accurate sensing model based on Maxwell's equations to expand the boundaries of Wi-Fi material sensing.

3 CONSTRUCTION AND SOLUTION OF ELECTRIC FIELD SCATTERING MODEL

The dielectric properties vary with the material (*i.e.* different materials have different complex permittivity), which makes them uniquely scatter incident waves [8, 18, 34]. As a result, the received signals will be affected by the material, shape, and position of target. To complete the image, we tried to parse the target's position, material, and shape information from the received signal.

Unlike the ray-tracing model, we construct the signal scattering model directly from the Maxwell equations, and then solve the dielectric distribution in the sensing domain.

3.1 Electric Field Scattering Sensing Model

Similar to much previous excellent work [25, 33, 44, 46, 52], *we consider the propagation of transverse-magnetic (TM) waves in 2-D space.* Specifically, as shown in Fig. 3, the electric field is perpendicular to the propagation direction of the wave, and the plane where the sensing domain D is located is parallel to the propagation direction of the wave. Similar scenarios are common. For example, we can use this model to describe when we place the receiving and transmitting antennas vertically on the table to sense the target on the table.

The key idea for building the model comes from the fact that the scattered field we receive can be regarded as the radiation field of the equivalent current in the sensing domain. The equivalent current depends on **the distribution of the complex permittivity**. The model used by *LiquImager* reflects the interaction between the scatterers within the scatterer and is closer to the real situation. We first introduce the integral form of the model and then discretize it for computational convenience.

The basic setup of the model is as follows. (1) There are P transmit antennas, where the i-th antenna is located at $\mathbf{r}_{txi}$. (2) There are Q receiving antennas, where the i-th antenna is located at $\mathbf{r}_{rxi}$. (3) The scattering domain D are discretized into $M \times M$ subunits, and the center of the i-th subunits is located at $\mathbf{r}_{di}$. (4) The liquid is placed in the domain D, and the relative complex permittivity of the subunits at position $\mathbf{r}_{di}$ is $\epsilon(\mathbf{r}_{di})$. The area outside the container is air, whose relative complex permittivity is $\epsilon(\mathbf{r}) = \epsilon_0$.

According to the principle of electric field superposition [2, 8, 18], *the total field can be regarded as the superposition of the incident field and the scattered field.* As a result, the total electric field $E_t(\mathbf{r})$ at position $\mathbf{r}$ is given by

$$E_t(\mathbf{r}) = E_i(\mathbf{r}) + E_s(\mathbf{r}), \tag{3}$$

where $E_i(\mathbf{r})$ and $E_s(\mathbf{r})$ denote the incident and scattered fields, respectively.

According to Eq. 1, *the scattered wave can be equivalent to the integral of the equivalent current radiation in the scattering domain D.* For TM waves in 2-D space, the equations are given by:

$$\begin{cases} E_t(\mathbf{r}) = E_i(\mathbf{r}) + k_0^2 \int_D G(\mathbf{r}, \mathbf{r}') I(\mathbf{r}') d\mathbf{r}' & \text{for } \mathbf{r} \in D \\ E_s(\mathbf{r}) = k_0^2 \int_D G(\mathbf{r}, \mathbf{r}') I(\mathbf{r}') d\mathbf{r}' & \text{otherwise,} \end{cases} \tag{4}$$

where k_0 is the wavenumber of the air. $G(\mathbf{r}, \mathbf{r}') = -\frac{j}{4} H_0^2(k_0|\mathbf{r} - \mathbf{r}'|)$ is the 2-D free space Green's function, where $H_0^2(.)$ is the 0-th order Hankel function of the second kind and $j^2 = -1$. The equivalent current density $I(\mathbf{r})$ is $I(\mathbf{r}) = [\epsilon(\mathbf{r}) - 1] E_t(\mathbf{r})$. Since the equivalent current pleases the distribution of the complex permittivity, combined with Eq. 3 and Eq. 4, we get the exact relationship between the received signal and the complex permittivity.

For ease of calculation, we discretize the equations [31]. In the domain D, the total field $\mathbf{E}_t$ is given by

$$\mathbf{E}_t = \mathbf{E}_i + \mathbf{G}_D \mathbf{I} \tag{5}$$

where $\mathbf{E}_t$, $\mathbf{E}_i$, and $\mathbf{I}$ are M^2 dimensions vectors, whose i-th elements are $\mathbf{E}_t(i) = E_t(\mathbf{r}_{di})$, $\mathbf{E}_i(i) = E_i(\mathbf{r}_{di})$, and $\mathbf{I}(i) = I(\mathbf{r}_{di})$, respectively. The equivalent current density is given by

$$\mathbf{I} = \mathbf{\Lambda}\mathbf{E}_t, \tag{6}$$

where $\mathbf{\Lambda}$ is a diagonal matrix with $\mathbf{\Lambda}(i,i) = \epsilon(\mathbf{r}_{di}) - 1$. The Green's coefficient matrix $\mathbf{G}_D$ is $M^2 \times M^2$ dimensions with $\mathbf{G}_D(m,n) = k_0^2 \iint_{D_n} G(\mathbf{r}_{dm}, \mathbf{r})d\mathbf{r}$, where D_n is the n-th subunits. Similarly, at the Q receiving antennas, scattered field is

$$\mathbf{E}_s = \mathbf{G}_S\mathbf{I} \tag{7}$$

where $\mathbf{E}_s$ is a Q dimensions vector, whose i-th elements are $\mathbf{E}_s(i) = E_s(\mathbf{r}_{rxi})$. The Green's coefficient matrix $\mathbf{G}_S$ is $Q \times M^2$ dimensions with $\mathbf{G}_S(q,n) = k_0^2 \iint_{Dn} G(\mathbf{r}_{rxq}, \mathbf{r})d\mathbf{r}$, where D_n is the n-th subunits of domain D.

Eq. 5 to Eq. 7 accurately describe the scattering of TM waves by liquid. Since different media (including air and various liquids) have different complex permittivity, liquid identification and imaging will become possible if we can obtain the complex permittivity in each discretized subunits by solving the model.

3.2 Obtaining the Permittivity Distribution with Backpropagation Scheme

We first consider a simple case: the scattered field $\mathbf{E}_s$ and the incident field $\mathbf{E}_i$ of the sensing domain D are known. In this subsection, we introduce how to estimate the distribution of the complex permittivity with this assume. Then we will generalize it to specific Wi-Fi perception tasks (Sec. 5).

It is difficult to solve this equation directly. The difficulty comes from two aspects. On the one hand, the Green's operator ($\mathbf{G}_S$ and $\mathbf{G}_D$) has filtering properties, which makes the original problem ill-conditioned [2]. On the other hand, the multiple scattering effect of the wave in the medium makes the problem strongly nonlinear.

Therefore, *we use the backpropagation scheme to calculate the dielectric distribution,* which can avoid calculations such as matrix inversion and singular value decomposition, and is applicable to the solution of any incident field, near field and far field [2]. As shown in Eq. 7, the scattered field $\mathbf{E}_s$ can be regarded as a function of the equivalent current $\mathbf{I}$. We first use the least squares method to estimate the equivalent current with the scattered field. The distribution of the complex permittivity is then estimated similarly. The steps are as follows.

(1) *Eliminate the equivalent current.* We assume that the equivalent current $\mathbf{I}$ in the scattering domain is linear with the scattering field, *i.e.*

$$\tilde{\mathbf{I}} = \xi\mathbf{G}_S^{\mathrm{H}}\mathbf{E}_s \tag{8}$$

where $\mathbf{G}_S^{\mathrm{H}}$ is the conjugate transpose of $\mathbf{G}_S$. The coefficient ξ can be obtained by the method of least squares, which is given by

$$\xi = \arg\min_{\xi} \|\mathbf{E}_s - \mathbf{G}_S\tilde{\mathbf{I}}\| = \frac{(\mathbf{E}_s)^{\mathrm{T}}(\mathbf{G}_S(\mathbf{G}_S^{\mathrm{H}}\mathbf{E}_s))^*}{\|\mathbf{G}_S(\mathbf{G}_S^{\mathrm{H}}\mathbf{E}_s)\|^2}, \tag{9}$$

where A^{T} and A^* are the transpose and conjugate of matrix A, respectively.

(2) *Eliminate the distribution of the complex permittivity.* Substituting Eq. 8 and Eq. 9 into Eq. 5, the total field is estimated to be

$$\tilde{\mathbf{E}}_t = \mathbf{E}_i + \mathbf{G}_D\tilde{\mathbf{I}} = \mathbf{E}_i + \xi\mathbf{G}_D\mathbf{G}_S^{\mathrm{H}}\mathbf{E}_s \tag{10}$$

With the least squares method and Eq. 6, an analytical solution can be obtained for the complex permittivity $\epsilon(\mathbf{r}_{di})$ of the i-th subunits with

$$\epsilon(\mathbf{r}_{di}) - 1 = \mathbf{\Lambda}(i) = \frac{\sum_{p=1}^P \tilde{\mathbf{I}}^p(i)\left[\tilde{\mathbf{E}}_t^p(i)\right]^*}{\sum_{p=1}^P \|\tilde{\mathbf{E}}_t^p(i)\|^2}, \tag{11}$$

In summary, imaging with backpropagation scheme of two main steps:
- Obtain the scattered electric field $\mathbf{E}_s$ at the receiving antennas and the incident electric field $\mathbf{E}_i$ at the scattering domain D;
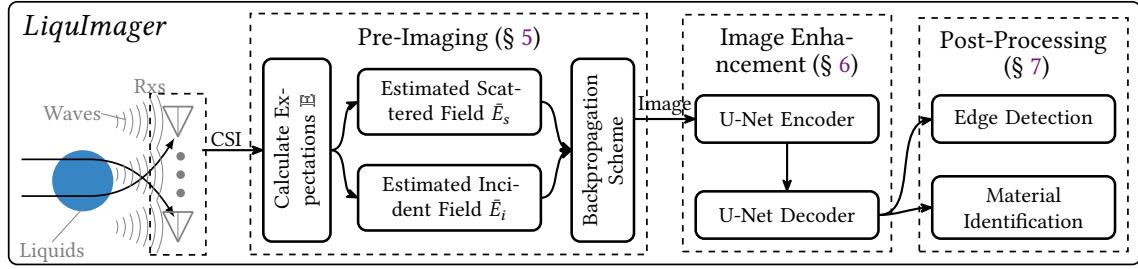
Fig. 4. Overview of the *LiquImager*. We first solve the electric field scattering sensing model to obtain the distribution of the complex permittivity (Sec 5). *LiqU-Net* is then used to improve the image on the pre-imaging results (Sec 6). Finally, we utilize the enhanced results for object imaging and liquid identification (Sec. 7).

- According to Eq. 8, Eq. 9, and Eq. 5, the equivalent current $\tilde{\mathbf{I}}$ and total field $\tilde{\mathbf{E}}_t$ at the domain D are estimated, and then the complex permittivity is obtained by using Eq. 11.

Once we have obtained the complex dielectric coefficient for each lattice point, we can complete the imaging.

4 OVERVIEW

Thanks to the fact that our model is directly based on Maxwell's equations, it makes it possible to carefully describe the effect of container position, shape, and liquid material on the received signal. The backpropagation scheme (Sec. 3.2) makes it possible for target imaging. However, the data required to solve the model is difficult to obtain directly. In addition, CSI data collected by commercial Wi-Fi devices has a lot of noise, which further increases the difficulty of imaging. In order to utilize CSI data for position-independent liquid identification and imaging, we designed *LiquImager*. The system diagram is shown in Fig. 4. First, we use the CSI data obtained by the receiving antenna for pre-imaging, and then use a deep learning network for image enhancement. Finally, we use the enhanced results for liquid identification and imaging. Specifically, it includes three components:

- **Pre-imaging**: We use the CSI data obtained by the receiving antenna to reconstruct the image of the liquid.
- **Image enhancement**: We build a deep learning network to enhance the image reconstructed by pre-imaging.
- **Post-processing**: We use the enhanced image to identify the liquid and the shape of the container.

5 PRE-IMAGING

To pre-image utilizing the backpropagation scheme (Sec. 3.2), we need to acquire the scattering field $\mathbf{E}_s$ at the receiving antenna and the total field $\mathbf{E}_t$ at the scattering domain D. However, they are difficult to measure directly. The only data we can get is the CSI data provided by commercial Wi-Fi devices, which can be regarded as the total field $\mathbf{E}_t^r$ at the receiving antenna. Therefore, we first introduce how to use the total field $\mathbf{E}_t^r$ at the receiving antenna to estimate the scattering field $\mathbf{E}_s$ at the receiving antenna and the total field $\mathbf{E}_t$ of the scattering domain D. Unlike ideal electric field data, CSI data contains a lot of noise [10, 47, 56]. Therefore, we modified the scheme to make it suitable for CSI data. For the convenience of expression, we only use the data of one frequency as an example to introduce the solution process, which is applicable to all sub-carriers.

In summary, it has three main steps:

- *Initialization.* We first manually determine the positions of the receiving antenna and transmitting antenna, and collect a set of data (denoted as E_i) when there is no target in the scattering domain.
- *Discretization.* (1) After artificially dividing the sensing domain, discretize it and obtain some intermediate variables needed in subsequent calculations, such as $\mathbf{G}_D$ and $\mathbf{G}_S$. (2) We use the received signal and distance
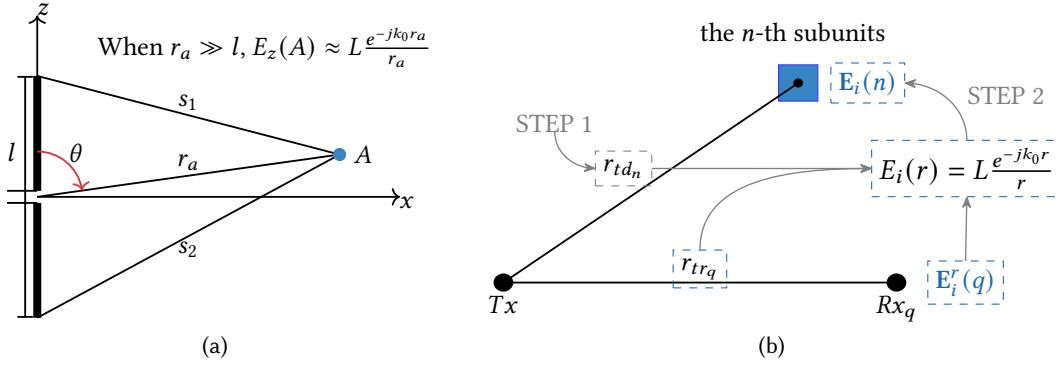
Fig. 5. Estimate the incident field for the sensing domain. (a) The far field of a transmitting antenna as a function of distance. (b) For the n-th subunits in the sensing domain D, we first estimate its distance r_{td_n} from the transmitter, and then combine the incident field at the receiving antenna $\mathbf{E}_i^r(q)$ to estimate its incident field $\mathbf{E}_i(n)$.

to estimate the two variables required to solve the model: the scattered wave E_s and the incident wave E_i in the sensing area, which are shown in Sec. 5.1 and Sec. 5.2, respectively.

- *BP solution.* In the presence of scatterers, we collect signals. The distribution of the complex permittivity in the scattering region is then estimated using a backpropagation algorithm, which is described in Sec. 5.3.

5.1 How to Estimate the Scattering Field $\mathbf{E}_s$?

Thanks to the superposition nature of electric fields [2, 8, 18], *we can obtain the scattering field by calculating the difference between the electric field in presence and absence.* In particular, when the target is not present, since the absence of scattering, the signal $\mathbf{E}_{t,w/o}^r$ at the receiving antenna contains only the radiated field of the transmitting antenna. When the target is present, the total field $\mathbf{E}_{t,w}^r$ is the superposition of the radiation field of the transmitting antenna and the scattering field $\mathbf{E}_s$. We use the difference method to estimate the scattered field, which is given by

$$\mathbf{E}_s = \mathbf{E}_{t,w}^r - \mathbf{E}_{t,w/o}^r. \tag{12}$$

5.2 How to Estimate the Incident Field $\mathbf{E}_i$?

The incident field $\mathbf{E}_i$ at domain D is more difficult to assess than the scattering data $\mathbf{E}_s$ because it is difficult to measure the electric field in that domain. Fortunately, the incident wave is a function of the transmission distance, which gives us the opportunity to reverse the incident field at domain D using the incident field at the receiving antenna.

5.2.1 The far field of a transmitting antenna is a function of distance. As shown in Fig. 5a, a thin straight antenna of length l parallel to the z-axis is placed at the origin. For any point A in the plane, the distance from the center of the antenna is r_a, and the distance from the upper and lower ends of the antenna is s_1 and s_2, respectively. When the distance of A from the antenna is much greater than the length of the antenna (*i.e.* $r_a \gg l$), *the radiated field can be approximated as a function of distance.* The length of the Wi-Fi antenna is usually less than $0.2\,\mathrm{m}$ and the distance between the antenna and the target is usually greater than $1\,\mathrm{m}$. In this case, $s_1^2 = (l/2)^2 + r_a^2 - lr_a cos\theta \approx r_a^2$, where θ is the angle between point A and z-axes. Similarly, we can get $s_1 \approx r_a \approx s_2$. In addition, with the setting

of the transmitting device unchanged (*i.e.* without changing the transmitting antenna and transmit power), the length l, impedance Z, and current I_0 of the antenna can be constant. The z component of the electric field at point A can be approximately written as [24, 28]

$$E_z(A) = L\frac{e^{-jk_0r_a}}{r_a} \tag{13}$$

where I_0 is the maximum current of the antenna, Z is the impedance of the antenna, k_0 is the wave number. L is a variable related to wave number k_0 and antenna length l, which is defined as $L = \frac{-jI_0Z}{2}(1 - cos\frac{k_0l}{2})$.

Now, *we try to estimate the incident wave in the scattering domain D utilizing the signal received by the receiving antenna.* As shown in Fig. 5b, there are one receiving antenna (Rx_q) and one transmitting antenna (TX). The distance between the transmitting antenna and the receiving antennas is r_{tr_q}. As demonstrated in Sec. 5.1, when a target is absent in the domain, the field of incidence at the receiving antenna is equal to the total field. Therefore, the incidence field $\mathbf{E}_i^r(q)$ at the receiving antenna can be measured. In addition, the distances r_{trx} between the transmitting antenna and the receiving antenna are easy for us to measure and estimate. At this point, *the question becomes how to estimate the incident wave $\mathbf{E}_i(n)$ at the n-th subunits when $\mathbf{E}_i^r(q)$ are known.*

5.2.2 Estimate the incident wave in the scattering domain. We first estimate the distance r_{td_n} between the n-th subunits and the transmitting antennas. We then combine the properties of the electric field excited by the transmitting antenna and the received signal to estimate the incident field $\mathbf{E}_i(n)$.

(1) *STEP 1: Estimate the distance r_{td_n} between the transmitting antenna and the n-th subunits.* We first roughly measure the distance between the transmitting and receiving antennas, and then utilize an optimization method to determine the relative positions of the transmitting and receiving antennas. We assume that the relative position of the sensing domain and the receiving antenna is known. The discretization of the sensing domain is performed artificially, which means that we can determine the position of each subunit through the position of the transmitting antenna.

Since the wavelength of a Wi-Fi signal is only few centimeters, inaccurate distance estimation introduces large phase errors. Therefore, we use several measurements to determine the relative position between the receiving antennas and the transmitting antennas in an optimal way. In particular, there are P transmitting antennas, whose positions are $[\mathbf{r}_{tx_1}, \cdots, \mathbf{r}_{tx_P}]$, respectively. There are Q receiving antennas, which are located at $[\mathbf{r}_{rx_1}, \cdots, \mathbf{r}_{rx_Q}]$, respectively. We first measure the distance between each pair of receiving and transmitting antennas $[r_{t_1r_1}, \cdots, r_{t_1r_Q}, \cdots, r_{t_Pr_Q}]$, and the distance between different transmitting antennas $[r_{t_1t_2}, \cdots, r_{t_1t_Q}, \cdots, r_{t_{Q-1}t_Q}]$. Then we optimize the error between the estimated distance and the measured distance to determine the position of the transmitting antenna, which is given by:

$$\underset{\mathbf{r}_{tx_1}, \cdots, \mathbf{r}_{tx_P}}{\arg\min} \left[\sum_p \sum_q \left(|\mathbf{r}_{tx_p} - \mathbf{r}_{rx_q}| - r_{t_pr_q} \right)^2 + \sum_i \sum_{i \neq j} \left(|\mathbf{r}_{tx_i} - \mathbf{r}_{tx_j}| - r_{t_it_j} \right)^2 \right] \tag{14}$$

Once we have the position of the transmitting antenna, we can directly enter the vector coordinate calculation to determine the distance of the transmitting antenna from each subunits of the sensing domain.

(2) *STEP 2: Estimate the incident field $\mathbf{E}_i(n)$.* Intuitively, we can substitute the distance between transmitter and sensing domain into Eq. 13 to solve for the incident field $\mathbf{E}_i(n)$. However, it is difficult for us to obtain the constant L. Instead, we use measurements from the receiving antennas to estimate $\mathbf{E}_i(n)$.

$$\begin{cases} |\mathbf{E}_i(n)| = \left|\mathbf{E}_i^r(p)\right| r_{tr_p}/r_{td_n} \\ \theta_i(n) = \theta_i^r(p) + k_0(r_{tr_p} - r_{td_n}) + 2m\pi, m \in \mathbb{Z}. \end{cases} \tag{15}$$

where $\mathbf{E}_n^r(p)$ is the received field at the p-th receiving antenna, $\theta_n^r(p)$ is the phase of $\mathbf{E}_i^r(p)$. Similarly, the phase of $\mathbf{E}_i(n)$ is $\theta_i(n)$. And $\mathbf{E}_i(n)$ is given by $\mathbf{E}_i(n) = |\mathbf{E}_i(n)|exp(j\theta_i(n))$.

Besides, noting that we have Q receiving antennas, we can obtain Q estimates. To improve estimation accuracy, we cluster them. Once both the scattered field $\mathbf{E}_s$ and the incident field $\mathbf{E}_i$ have been estimated, we can pre-image the sensing domain D utilizing the method described in Sec. 3.2.

5.3 How to Do Pre-imaging in the Presence of Noise?

In Sec. 5.1 and Sec. 5.2, we show how to use the total field data received by the receive antennas to estimate the scattered and incident fields for pre-imaging. Unfortunately, the CSI collected by commercial Wi-Fi devices contains a lot of noise, making it difficult for us to obtain an accurate total field. Besides, since we are imaging utilizing the values of the individual position fields rather than their relative values, lots of methods [7, 41] that work well but require the use of a reference antenna are difficult to apply. Fortunately, we noticed that when the environment (temperature, humidity, device state, *etc.*) does not change significantly, noise can be regarded as a generalized stationary random signal [39]. This gives us the opportunity to use statistical methods to make pre-imaging immune to signal noise. In the following, we first analyze the characteristics of Wi-Fi noise, and then introduce how we process the Wi-Fi signal so that the noise does not affect the imaging.

5.3.1 Multiplicative noise is difficult to remove. As mentioned in many excellent works [10, 47, 56], the data collected by commercial Wi-Fi equipment contains a lot of noise. *These noises can be divided into two categories:*

- *Multiplicative noise.* In the process of CSI estimation, due to the inherent error of the hardware of the receiving device, the amplitude and phase of the received data will be shifted. For example, automatic gain control (AGC) scales the amplitude of ideal data. And the carrier frequency offset (CFO) caused by inaccurate clock synchronization and the like will cause the phase of the received data to deviate from the ideal data. These shift can be equivalent to the product of a complex number and ideal data.
- *Additive noise.* During the transmission process, the signal will be affected by factors such as equipment thermal noise, which is independent of the signals. The final received CSI appears as a linear superposition of the received data and these noises.

As a result, with one transmitting antenna, the CSI received by the p-th antenna $\hat{\mathbf{E}}_t^r(p)$ can be expressed as:

$$\hat{\mathbf{E}}_t^r(p) = N_{m,p}(t) \cdot \mathbf{E}_t^r(p) + N_p(t). \tag{16}$$

We use $N_{m,p} = A_p(t)e^{j\phi_p(t)}$ to denote the multiplicative noise. $N_p(t)$ is the additive noise. $A_p(t)$ and $\phi_p(t)$ are the amplitude and phase shift caused by multiplicative noise at time t, respectively. $\mathbf{E}_t^r(p)$ is the ideal CSI. And t is the time index. For additive noise, we usually regard it as Gaussian noise with a mean of 0, *i.e.* $N_p(t) \sim \mathcal{N}(0, \sigma^2)$, where σ^2 is the variance of the noise [36]. For multiplicative noise, there are many works [7, 41] that have carried out thorough research, which show that in a multiple-input and multiple-output (MIMO) Wi-Fi system, the multiplicative noise of each receiving antenna at the same time is the same, *i.e.* $N_{m,1}(t) = N_{m,2}(t) \cdots = N_{m,Q}(t) = N_m(t)$, where Q is the number of receiving antennas.

*Although there are many multiplicative noise cancellation methods (*e.g. *ratio method) based on multi-antenna systems, they are no longer applicable* since *LiquImager* uses data from both the incident field and the scattered field. Those methods take advantage of properties exhibited by Eq. 16. Specifically, they often choose a receiving antenna as a reference, and then utilize the ratio of the data from other antennas to the reference antenna data to complete the sensing task. However, For *LiquImager*, the ratio data is difficult to apply. As introduced in Sec 5.1, we utilize the difference in the electric field in the presence and absence of the target to estimate the value of the scattered field, *i.e.* $\mathbf{E}_s^r = \mathbf{E}_{t,w}^r - \mathbf{E}_{t,w/o}^r$. Since it is difficult for us to estimate the relationship between the data ($E_{t,w}^{ref}$ and $E_{t,w/o}^{ref}$) received by the reference antenna in these two cases, it is difficult to utilize the difference of the reference values $\left(\mathbf{E}_{t,w}^r / E_{t,w}^{ref} - \mathbf{E}_{t,w/o}^r / E_{t,w/o}^{ref} \right)$ to estimate the scattered field.
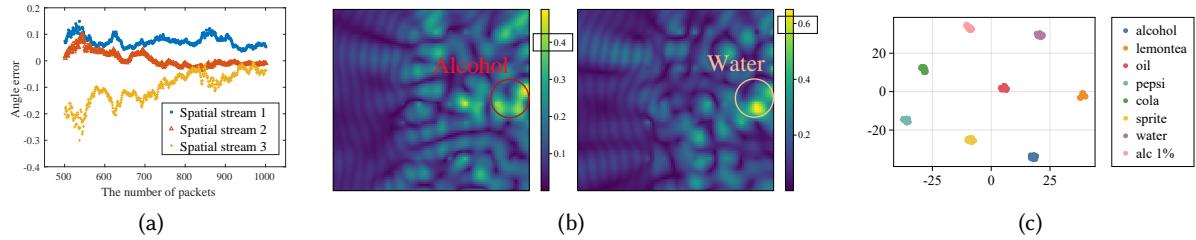
Fig. 6. Visualization. (a) When the number of data packets expected to be used to estimate the signal exceeds 700, the error between the phase difference between different receiving antennas and the result obtained by the ratio method is less than 0.1. (b) After preprocessing the signals collected at the same location, water and alcohol exhibit different characteristics. (c) For 20 independent collections of data, there are large differences between different liquid characteristics.

5.3.2 Using statistical methods to insulate imaging from noise. Multiplicative noise comes from equipment-related factors such as automatic gain control (AGC) and antenna gain. When the environment does not change drastically, we believe that its expectation is independent of time [39], *i.e.* $\mathbb{E}[N_m(t)] = \mathfrak{R}_m$, where [.] is the expectation operator. Therefore, in the case of constant environment, the expectation of the received signal is given by:

$$\mathbb{E}[\hat{\mathbf{E}}_t^r(p)] = \mathfrak{R}_m \cdot \mathbf{E}_t^r(p) + \mathbb{E}[N_p(t)] = \mathfrak{R}_m \cdot \mathbf{E}_t^r(p). \tag{17}$$

If we use the expectation of CSI to estimate the scattered and incident fields, the results of the pre-imaging are not affected by the noise. Specifically, Eq. 12 (for the estimate of the scattered field) and Eq. 15 (for the estimated incident field) are rewritten as

$$\begin{cases} \hat{\mathbf{E}}_s = \mathbb{E}\left[\hat{\mathbf{E}}_{t,w}^r\right] - \mathbb{E}\left[\hat{\mathbf{E}}_{t,w/o}^r\right] = \mathfrak{R}_m \cdot \mathbf{E}_s. \\ \hat{\mathbf{E}}_i = \mathfrak{R}_m \cdot \mathbf{E}_i. \end{cases} \tag{18}$$

Further, we use the estimated value of the electric field at this time for pre-imaging (Sec. 3.2), and the result is not affected by noise. In particular, Eq. 8, Eq. 9, and Eq. 10 are rewritten as:

$$\begin{cases} \hat{\xi} = \dfrac{(\hat{\mathbf{E}}_s)^{\mathrm{T}}(\mathbf{G}_S(\mathbf{G}_S^{\mathrm{H}}\hat{\mathbf{E}}_s))^*}{\|\mathbf{G}_S(\mathbf{G}_S^{\mathrm{H}}\hat{\mathbf{E}}_s)\|^2} = \dfrac{\mathfrak{R}_m^2(\mathbf{E}_s)^{\mathrm{T}}(\mathbf{G}_S(\mathbf{G}_S^{\mathrm{H}}\mathbf{E}_s))^*}{\mathfrak{R}_m^2\|\mathbf{G}_S(\mathbf{G}_S^{\mathrm{H}}\mathbf{E}_s)\|^2} = \xi \\ \hat{\tilde{\mathbf{I}}} = \hat{\xi}\mathbf{G}_S^{\mathrm{H}}\hat{\mathbf{E}}_s = \xi\,\mathfrak{R}_m\mathbf{G}_S^{\mathrm{H}}\mathbf{E}_s \\ \hat{\mathbf{E}}_t = \hat{\mathbf{E}}_i + \mathbf{G}_D\hat{\mathbf{I}} = \mathfrak{R}_m\mathbf{E}_t \end{cases} \tag{19}$$

Therefore, with Eu. 11 the complex permittivity of the *i*-th subunits is estimated as

$$\hat{\epsilon}(\mathbf{r}_{di}) - 1 = \frac{\sum_{p=1}^{P}\hat{\tilde{\mathbf{I}}}^p(i)\left[\hat{\tilde{\mathbf{E}}}_t^p(i)\right]^*}{\sum_{p=1}^{P}\|\hat{\tilde{\mathbf{E}}}_t^p(i)\|^2} = \frac{\mathfrak{R}_m^2\sum_{p=1}^{P}\tilde{\mathbf{I}}^p(i)\left[\tilde{\mathbf{E}}_t^p(i)\right]^*}{\mathfrak{R}_m^2\sum_{p=1}^{P}\|\tilde{\mathbf{E}}_t^p(i)\|^2} = \frac{\sum_{p=1}^{P}\tilde{\mathbf{I}}^p(i)\left[\tilde{\mathbf{E}}_t^p(i)\right]^*}{\sum_{p=1}^{P}\|\tilde{\mathbf{E}}_t^p(i)\|^2}, \tag{20}$$

which is equal to the result without noise.

Fig. 6 shows the results of some confirmatory experiments. Since the expectation of a random signal can be estimated by the average of multiple measurements, we first test how many packets the mean can better characterize the expectation of the signal. Specifically, we use the ratio method to obtain the phase difference between the two receive antennas as the ground truth. Then we take a certain number of packets from the original received data to take the average. Those average values are then used to estimate the phase difference

between the different antennas. Fig. 6a presents the results on data from different spatial streams, and we find that the error is less than 0.1 radians when the number of packets used to estimate the expectation exceeds 700.

We collect data at the same location (*A*). For 8 kinds of liquids, data were collected independently 20 times. For each independent collection, we place the container at a random collection point around *A*. The distance difference between the collection point and *A* is less than $\sqrt{2}$ cm.

Fig. 6b shows the results after data preprocessing. For two liquids, alcohol and water, their pretreatment results show different characteristics. For example, they have different brightness ranges, the left image has a brightness range less than 0.5, while the right image has a brightness range greater than 0.6. Furthermore, their "patterns" are inconsistent, caused by their different scattering of the signal. We simply flatten the preprocessing results into a one-dimensional vector as the characteristic of the liquid, and then use the t-SNE algorithm for dimensionality reduction. Fig. 6c shows the results after dimensionality reduction. *We find that the degree of differentiation between different liquid characteristics is relatively large.* This provides us with the opportunity to conduct subsequent multi-position material identification. But we also noticed that it is difficult to perfectly restore the shape of the container. We believe this is because COTS Wi-Fi devices are difficult to accurately measure electric field data, so we design an image augmentation network to improve the pre-imaging results.

6 IMAGE ENHANCEMENT USING *LIQU-NET*

Our goal is to recover the shape, position, and material of targets from low-resolution images which are given by the pre-imaging module. A typical approach is to first encode an image with a convolutional neural network and then decode it to obtain a high-resolution image. We note that our task is similar to the goal of semantic segmentation, so based on U-Net [32], a classic method in semantic segmentation, we design the *LiqU-Net* for encoding-decoding.

6.1 Problem Analysis and Encoder-decoder Selection

If material identifying is not considered, the goal of our image augmentation module is similar to binary semantic segmentation. Specifically, we want to classify the subunits in the image where the target is located into one class and the subunits where no target exist into another class. Many previous studies [1, 17, 55] have shown that the U-Net network can achieve good results when dealing with semantic segmentation tasks, and it can be used to solve equations [21, 42]. Based on U-Net, we introduce a network called *LiqU-Net* for image enhancement.

However, due to the differences in the input and output of the network, there are two practical issues that need to be solved first when applying the U network to target domain imaging.

- *Input.* The input of the U-Net network is 3-channel RGB data, but the result of our pre-imaging is the single-channel complex permittivity. We notice that directly using complex numbers as the input of the deep learning network will bring huge storage and computing overhead, and many mainstream deep learning frameworks do not support networks composed of complex numbers well. Therefore, the first practical issue that needs to be solved is what kind of data is used as the input of the *LiqU-Net* network.
- *Output.* The output of the U-Net network is a binary geometry (for example, only two values of 0 and 1), but this cannot complete the material identification task. The second practical problem to be solved is how to add material identification to the *LiqU-Net* network.

Therefore, we modify the LiqU-Net *network to adopt it to our image enhancement by combining the characteristics of electromagnetic signals.*

Input. (1) Instead of directly feeding the permittivity into the *LiqU-Net* network, we feed the imaginary and amplitude of the complex permittivity as two separate channels. The reason is as follows. On the one hand, since many materials have similar dielectric properties, we expect more dimensional data for material classification. Both the real and imaginary parts of the complex permittivity are determined by the polarization properties of
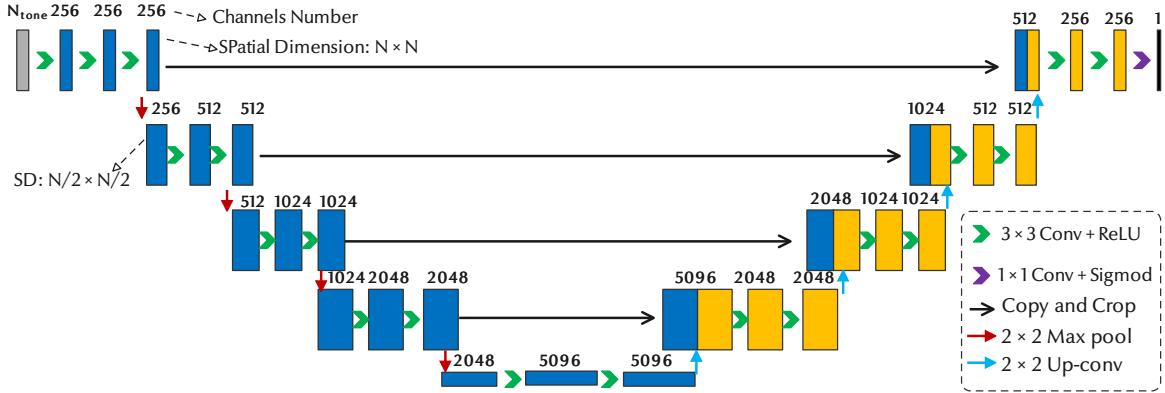
Fig. 7. The structure of the *LiqU-Net*. We first encode the image using a 4-layer network (left), and then use a 4-layer decoder to enhance the image (right).

the material itself, which means that both can be used as unique features of the material. Therefore, theoretically, the segmentation of the target domain can be completed no matter whether the data of the real part or the imaginary part of the complex permittivity is used. However, a portion of the complex permittivity of many materials is similar. For example, the complex permittivity of Coca-Cola and Sweet Tea are $71.25 + 18.87j$ and $71.23 + 16.72j$, respectively [5]. The difference in imaginary parts (2.15) is more pronounced than the difference in their real parts (0.02). Therefore, multichannel data helps to improve the accuracy of material identification. On the other hand, nodes in the network may not be activated because sometimes the imaginary part data will be very close together. For many lossless media, the imaginary part of their complex permittivity is 0 [20]. For example, the complex permittivity of Motor Oil is $2 + 0j$ [5]. Therefore, we do not directly input the imaginary part of the complex permittivity into the *LiqU-Net* network, but utilize its amplitude to embody the imaginary part information. (2) We feed data from multiple subcarriers as independent channels. The methods presented in Sec. 5.3 are pre-imaging for single-frequency signals. However, commercial Wi-Fi equipment can collect CSI data of multiple frequencies, and we use them as different channels of input, which can suppress the influence of noise.

Output. In order to complete the material identification, an intuitive way is to take the complex permittivity of the target domain as the output. However, actually accurate permittivity measurements are expensive. Therefore, we use different intervals in $[0, 255]$ to represent different materials. Specifically, if our goal is to distinguish T kinds of materials, we first divide $[0, 255]$ into T different intervals, and then use the center of each interval as the true value of the material for supervised training.

6.2 Network Structure

In Fig. 7, we illustrate the architecture of *LiqU-Net*. The specific details of the network are as follows.

(1) *Input.* The input of the network is a three-dimensional tensor whose shape is $[N_{tone} \times 2, M, M]$. Among them, N_{tone} is the number of subcarriers, and M is the number of grids discretized in the scattering domain.

(2) *Convolutional network.* The network consists of an encoding part (left) and a decoding part (right). The encoding part consists of repeated application of 3×3 (the padding is 0 and stride is 1) convolution and rectified linear unit (ReLU). Then there is a 2×2 maximum pooling operation, which can improve the receptive field of the network. The structure of the decoding part is similar to that of the encoding part, except that the 2×2

maximum pooling operation is replaced by a 2×2 transposed convolution, which is used to restore the size of the image. Besides, we insert skip connections before the encoder and decoder. This allows more dimensionality and location information to be preserved in the network. Therefore, the latter layer can freely choose the features in the shallow layer or that in the deep layer, which is more advantageous for semantic segmentation tasks. Furthermore, it alleviates the vanishing gradient problem during training [12, 21].

(3) *Output*. The output of the network is a two-dimensional tensor whose size is $[M, M]$. Among them, M is the number of grids discretized in the scattering domain. Different values are used to represent different materials.

(4) *Loss function*. We have two goals: on the one hand, we hope that *LiqU-Net*'s segmentation of the target domain is accurate, which is conducive to judging the shape and position of the target; on the other hand, we hope that the output value of the target domain is the same as the ground truth, which is helpful for material identification. Therefore, we use the sum of the binary-cross-entropy loss and the mean square error loss as the total loss function, which is defined as follows:

$$\mathcal{L} = \mathcal{L}_{bce}\left(\frac{\max \hat{y} - \hat{y}}{\max \hat{y} - \min \hat{y}}, \frac{\max y - \hat{y}}{\max y - \min y}\right) + \mathcal{L}_{mse}(\hat{y}, y), \tag{21}$$

where $\mathcal{L}_{bce}$ and $\mathcal{L}_{mse}$ are the binary-cross-entropy loss and the mean square error loss, respectively. $\hat{y}$ and y are the output of the network and the ground truth, respectively. The maximum and minimum values of $\hat{y}$ are $\max \hat{y}$ and $\min \hat{y}$, respectively.

7 CONTAINER IMAGING AND LIQUID IDENTIFICATION

After obtaining the complex permittivity distribution of the sensing domain D, container imaging and liquid identification can be accomplished with only simple processing. Recall that in *LiqU-Net*'s output, different complex permittivity maps to different pixel values, which means that different colors in the image correspond to different materials. For example, if we use blue to represent water and black to represent air, then when there is a cylindrical container that is full of water in the sensing domain, a blue circle will appear in the position corresponding to the container in the output, and the other areas will be black. Therefore, we only need to perform edge detection to determine the outline of the container, and material identification can be done by judging the size of the pixel value of the area where the liquid is located.

7.1 Determining Container Shape by Edge Detection

The edge detection scheme of the image is often used to distinguish the target from the background. It is based on a basic assumption that the difference between the pixels inside the target domain and that inside the background domain is relatively small, but the pixel value changes significantly in the transition domain from the target to the background. In the edge part, the pixel value has a "jump" or a large change. As a result, the first-order derivative of the pixels on the edge will have an extreme value, and the second-order derivative will be zeros. Based on this principle, edge detection can be performed.

(1) We use *Canny* operator for edge detection. We first use Gaussian filtering to filter the possible noise in the point image.

(2) Then we use the *Canny* operator to calculate the gradient $U(x, y)$ and angle $\theta_u(x, y)$ of the image, which are given by:

$$U(x, y) = \sqrt{U_x^2(x, y) + U_y^2(x, y)},$$
$$\theta_u(x, y) = \arctan\left(\frac{U_y(x, y)}{U_x(x, y)}\right), \tag{22}$$
$$U_x = \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix}, \quad U_y = \begin{bmatrix} -1 & -1 \\ 1 & -1 \end{bmatrix}$$
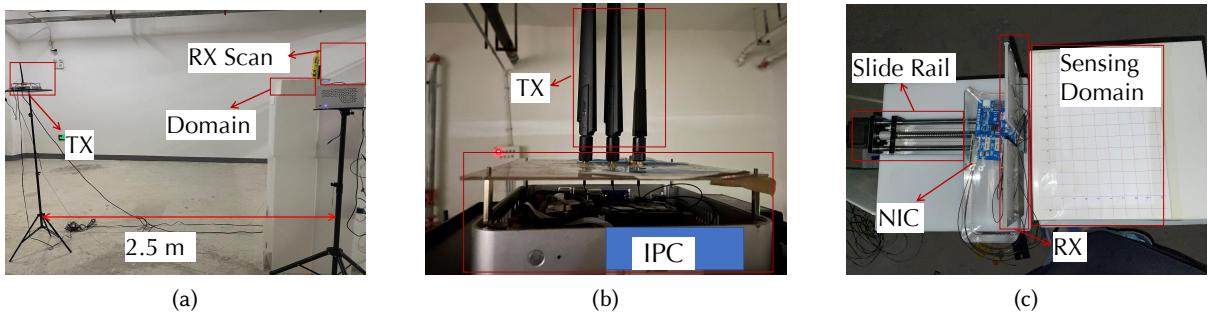
Fig. 8. The experiment setup of *LiquImager*. (a) The overall setup. (b) The transmitter setup. (c) The receiver setup.

where $U_x(x, y)$ and $U_y(x, y)$ are the horizontal and vertical gradients of the image, respectively.

(3) After we perform non-maximum suppression on the gradient image $U(x, y)$, we use the threshold method to determine the edge of the image. Specifically, we first compare each pixel with two neighboring pixels with the same gradient. If the gradient value is not the maximum value among the three pixels, the pixel is suppressed, that is, the pixel is zero. Then we judge the edge position according to the threshold.

7.2 Liquid Identification by K-Nearest Neighbor Classifier

Since the complex permittivity varies with the type of liquid, different types of liquid have different pixel values in the final imaging result. As described in Sec. 6, we set the pixel value corresponding to air to 0. As a result, We only need to extract the largest k value as the liquid feature from the imaging result, and then use the k-nearest neighbor (KNN) classifier to identify the liquid. Considering that in our setup, one pixel corresponds to an actual width of $25/100 \approx 0.25$ cm, and the width of the container is usually more than 4 cm, we set k to 15.

8 IMPLEMENT

8.1 Hardware

As shown in Fig. 8, we implement *LiquImager* utilizing commercial Wi-Fi NIC IWL5300, which supports three antennas. The NICs are installed on the computer with Intel Core i7-5550U CPU and 8 GB RAM. Specifically, we use one NIC for transmitting and two NICs for receiving. For the NIC that transmits the signal, we use three spatial streams, and for each receiver, we use three RF chains to receive the signal, *i.e.*, they form a 3×3 MI-MO system. And for the receiver, the antenna separation is 5 cm to avoid strong coherence. We place receiving antennas on two easy-to-move small handcarts. We utilize a computer with AMD 7950X cpu 64G memory to pre-imaging, and use NVIDIA 3090 graphics cards to train *LiqU-Net* networks for enhancing images.

8.2 Software and Data Collection

The part of *LiquImager* data collection, which is developed based on PicoScenes [19], running on Ubuntu 20.04 LTS. All algorithms in *LiquImager* are implemented by python 3.10 with CUDA 11.8 and Pytorch 2.0. All NICs work at 128 channels (*i.e.* , $5.755GHz/HT40-$). It is worth noting that work in different channels is feasible, because the complex permittivity of liquid is sufficiently different from that of air in the Wi-Fi band. To collect scatter data for more positions, we move the receiving antennas which is placed with receiving antennas with an electric slide rail (the length is 20 cm) to achieve virtual antenna array. Then, we control the sliding rail to move the receiving antennas 9 times, and the distance of each movement is 2 cm.
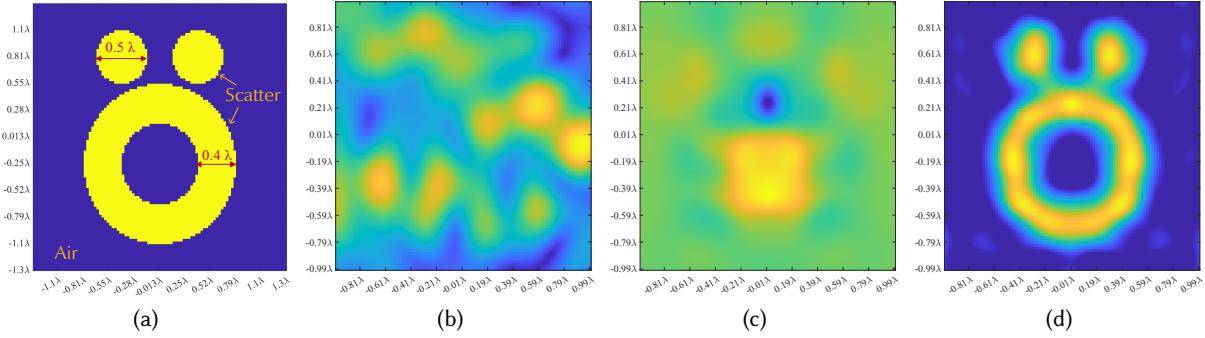
Fig. 9. Imaging results of ray tracing model and backpropagation model. (a) Schematic diagram of a scatterer. (b) Since the ray tracing scheme does not consider the interaction within the scatterer, when the size of the scatterer is small, it is only possible to vaguely identify the presence of scatterers in this area. (c) Imaging results of the pre-imaging module. (d) The imaging result after 50 iterations of the backpropagation algorithm.

8.3 Wi-Fi Devices Calibration

Most commercial Wi-Fi devices suffer from imperfect hardware implementations [3]. As a result, the collected signals are biased, which generally need to be removed before performing wireless sensing task.

(1) *Radio chain phase offset*. Every time the router is turned on, there is a random phase shift in the transmitter and receiver, which leads to inaccurate CSI measurements [48]. Fortunately, such offset is stable across all packets and Wi-Fi channels. As a result, we utilize coaxial cable to connect the transmitter (TX) and receivers (RX) in turn, and then collect the CSI measurements for each channel for get the phase offset of each TX-RX pair. Before performing any sensing task, we cancel such phase offset from the raw CSI data.

(2) *Amplitude outliers*. Due to the existence of automatic gain control (AGC) and the instability of the Wi-Fi device itself, even if the environment remains unchanged, there will be many outliers in the received signal. Therefore, before pre-imaging, we perform outlier detection on the received signal and use a Gaussian window to smooth the data.

9 REVISIT

In this section, we aim to illustrate *LiquImager* has better resolution than ray models and the respective roles of pre-imaging (Sec. 5) and *LiqU-Net* (Sec. 6) in the system through simulation verification.

9.1 Imaging Resolution

The resolution limit of the imaging scheme based on the ray tracing model is $\lambda/2$ [2], which is close to the container size. This is because the ray tracing model ignores the interaction between individual scattering points inside the scatterer. Since the field model used by *LiquImager* more accurately describes the impact of scatterers on signals, the resolution of using nonlinear methods to reversely infer the complex permittivity distribution in the sensing area can reach 0.13λ or even lower [4, 51]. The Wi-Fi signal frequency we use is 5.755 GHz, and its corresponding resolution limit is less than 1 cm, which is sufficient for identifying containers larger than 5 cm.

We performed simulation calculations for the two models. The result is shown in Fig. 9. We set the sensing area as a square with a side length of 2.5λ, which is shown in Fig. 9a. Set up a "rabbit" (yellow part in the figure) as a scatterer, surrounded by air (blue part). The relative complex permittivity of the scatterer is set to 2. And 16 transmitting antennas and 54 receiving antennas are placed at equal intervals on a circle 4λ away from the center

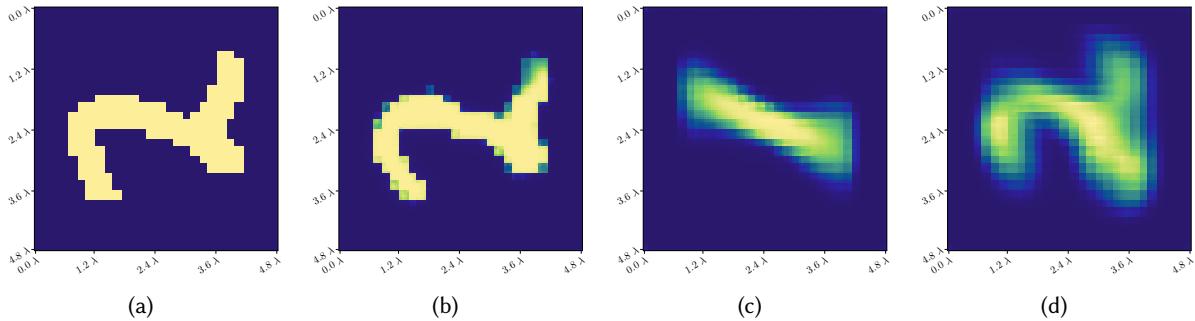|			
(a)	(b)	(c)	(d)

Fig. 10. It is easier to achieve good training results by first using the backpropagation algorithm for data preprocessing. (a) Schematic diagram of a scatterer. (b) Imaging results when first using the backpropagation algorithm for data preprocessing, and then using 20% of the training samples for training. (c) Imaging results when directly using 20% of the training samples for training. (d) Imaging results when directly using all training samples for training.

of the scattering area. Fig. 9b and Fig. 9d are image results reconstructed using ray tracing model [29] and field model, respectively. It can be found that in the imaging results of the ray tracing model, only the presence of scatterers in this area can be vaguely distinguished. But *in the imaging results of the field model, we can distinguish the basic shape and position of the scatterers.*

9.2 The respective roles of pre-imaging and *LiqU-Net*

In particular, due to the nonlinear fitting ability of the deep learning network, using *LiqU-Net* for radio frequency signal imaging *avoids the multiple iterations required by traditional solutions to the model*, which takes a long time (the iteration time may exceed 20s, but the neural network inference time is only 180ms). Besides, it can reduce the influence of ill-conditioned integral equations, thereby *improving imaging ability* [45, 49]. However, the acquisition and calibration of RF data are complex, and we lack samples with diverse shapes to construct a sufficiently rich sensing area. As a result, it is difficult to collect a sufficient number of RF data sets to directly train the deep learning model. So we borrowed image data sets to conduct relevant experimental verification. Using *physical models* for data preprocessing can help *reduce the amount of training required for network training* [42]. Our evaluation results show that *LiquImager* can achieve better imaging results by using only 20% of the data for training than directly training *LiqU-Net* with the entire data. The specific simulation verification is as follows.

(1) *Model-driven approaches often require multiple rounds of iterations to achieve better results when imaging strong scatterers* [2] (relative permittivity much greater than 1). Unfortunately, common liquids usually have large relative complex permittivity. For example, the relative complex permittivity of water is about 78+20j [5, 33]. Figure 9 shows the simulation results. For a scatterer with a relative complex permittivity of 2 (as shown in Figure 9a), the imaging algorithm of the backpropagation algorithm is shown in Figure 9c. The image is very blurry. In order to improve imaging clarity, many researchers try to optimize through iterative methods. Figure 9d shows the effect after 50 rounds of iterations. Although the image clarity is acceptable, the iterative algorithm takes a long time to execute (on a computer with an AMD 7950X processor and 64G memory, the time for 50 iterations is about *29s*). In addition, when the reflection intensity of the scatterer is greater, the imaging effect is still difficult to guarantee. However, *LiqU-Net* only takes about *180ms* to infer a result (on a computer with an AMD 7950X processor and 64G memory).

(2) Although deep learning networks help improve the performance of imaging using RF signals [15, 42, 45], purely data-driven solutions require a large amount of training data, which is difficult to collect. Compared
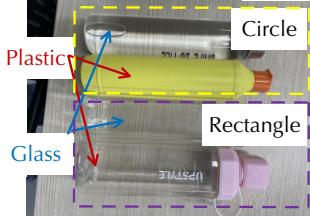
Fig. 11. The containers used in the experiment.

Table 1. Liquids tested and their corresponding colors

Liquid	Water (A)	Cola (B)	Alcohol (C)	Lemon tea drink (D)
Color	●	●	●	●
Liquid	Oil (E)	Pepsi (F)	Sprite (G)	Alcohol solution (1%) (H)
Color	●	●	●	●

with the data-driven approach, *the data-model dual-driven approach we adopt can achieve better results using less training data.* Figure 10 shows the results of simulation verification. For the convenience of data acquisition, we binarize the images in the classic handwriting data set MNIST and assign them different relative complex permittivity as scatterers. Specifically, for the image labeled i, we set the complex permittivity of the digital part to $i + 3$ and the remaining parts to 1. For example, Figure 10a shows an image labeled 2. We set the complex permittivity corresponding to the yellow area to $2 + 3 = 5$, and set the complex permittivity of the blue area to 1 to represent air. We use the data labeled 2 and 7 as the test set and the remaining data as the training set. For each image, we first calculate its scattering field using the method of moments. We first evaluated the effect of backpropagation and then training using *LiqU-Net*. The imaging effect of the scatterers in the test concentration is shown in Figure 10b. *We find that although only 20% of the samples in the training set were used for training,* LiqU-Net *still achieved good imaging results. It is worth noting that for the same scatterer with a complex permittivity of 5,* LiqU-Net *has stronger imaging capabilities than the model-driven iterative scheme.* However, when we use the scattered field data directly for training without using the backpropagation algorithm, *LiqU-Net*'s imaging effect is not good. When 20% of the data is used for training, *LiqU-Net* has difficulty completing the imaging task, and the results are shown in Figure 10c. Even though we use the entire training set for training, the imaging results are still blurry, which is shown in Figure 10d.

10 EVALUATION

We conduct a series of experiments to evaluate the performance of *LiquImager*. Firstly, we build a benchmark for fine-grained liquid identification with 8 different liquids and evaluate *LiquImager* on the benchmark. Then we validate the performance of *LiquImager* with different distances, and in containers with different materials and shapes. We then evaluate the performance of *LiquImager* for imaging three common container shapes. Finally, we explore the boundaries of *LiquImager*.

10.1 Dataset

We choose 8 different liquids, including water, Cola, alcohol, lemon tea drink, corn oil, Pepsi, Sprite, and 1% alcohol solution, to build the benchmark and evaluate the fine-grained classification performance of *LiquImager*. The types of liquids are shown in Table 1. Except for two alcohol-containing liquids, the others are purchased from supermarkets. Alcohol has an ethanol concentration of 99.9%, and the 1% alcohol solution is prepared using alcohol and water. To test *LiquImager*'s ability to identify liquids in different shapes and sizes, we selected four common containers with volumes ranging from 350 ml to 500 ml. In order to verify *LiquImager*'s imaging capability, we choose three common shapes of containers, including circle, rectangle, and rectangle rotated at a certain angle, and their materials include plastic and glass. The size of our sensing domain is 25 cm × 25 cm. In order to verify the location-independence, we selected 49 locations in the domain for data collection. In each collection location, we repeat the placement five times with slight and random displacement and orientation.
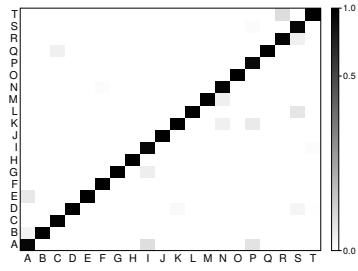
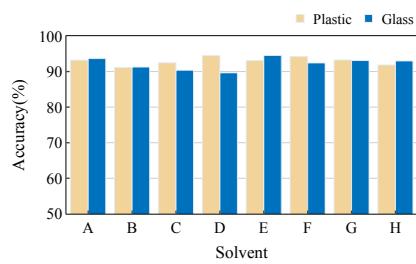Fig. 12. Identification performance for 20 liquids.

Fig. 13. Performance for different container materials.
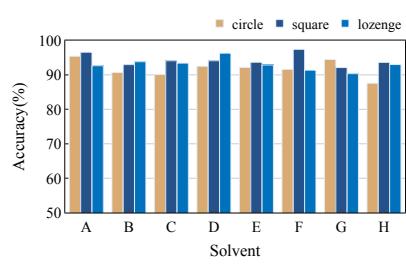
Fig. 14. Identification performance for different shapes.

The displacement range is within -1 cm to 1 cm. For the data at these five random *collection points*, we randomly selected one to join the test set and the remaining four to join the training set during five-fold cross validation. At each collection point, we collect approximately 1000 data packets (few Wi-Fi data packets will be lost during the actual collection process). Then we use a sliding window with a width of 900 to obtain 50 samples. In addition, in order to avoid the data obtained in this way being too approximate, we artificially add Gaussian noise with a variance of 0.2 to the completed data before it is input to *LiqU-Net* (added to both the training set and the test set). The temperature of all liquids is at room temperature between 16 °C to 20 °C. The accuracy results are calculated from the average of 5 results.

10.2 Liquids Identification

10.2.1 Basic experiment. We first evaluate *LiquImager*'s ability to identify liquids. We selected 20 different liquids, including (A) water, (B) cola, (C) alcohol, (D) lemon tea drink, (E) oil, (F) Pepsi, (G) sprite, (H) 1% alcohol solution, (I) Fanta, (J) soy sauce, (K) vinegar, (L) cooking wine, (M) orange juice (Huiyuan juice) , (N) apple juice (Huiyuan), (O) beer (Qingdao), (P) soju (Jinro), (Q) green tea beverage (Master Kong), (R) mango juice (Nongfu Spring), (S) milk (Deya), and (T) coffee (Starbucks). Some of these are similar liquids, such as Pepsi and Coke, water and 1% alcohol solution. They include several common liquid types, water, tea, coffee, carbonated drinks, condiments (soy sauce, etc.), juice and fat-soluble liquids (oil).

We use a plastic container with a diameter of 5 cm to hold them. For each liquid, we independently collected the signal 50 times. Having gotten the liquid feature, we can adopt a simple K-Nearest Neighbors algorithm (K=1) to differentiate those liquids. Then five-fold cross-validation is used to evaluate the recognition ability of the system. The results are shown in Fig. 12. The results show that for 20 liquids, the average recognition accuracy exceeds 90%.

10.2.2 Impact of container material. In real scenes, there are often more than one material for containers. We expect *LiquImager* to perform well for containers of different materials. In order to test the influence of different shapes of cups on liquid identification, we select four containers of different material for testing, whose materials include plastic and glass. The containers are shown in Fig. 11. For each container, we selected data from the test dataset with 10 different positions for evaluation. The data for each position contains 8 liquids. Having obtained liquids features, we can adopt a simple K-Nearest Neighbors algorithm (K=1) to differentiate those liquids.

The identification results are shown in Fig. 13. The results show that *LiquImager* achieves an accuracy over 90%, which means that *LiquImager* can identify liquids with high accuracy despite the material of the containers. We find that when the material of the container is different, the accuracy of liquid identification is not much different. We believe this is due to the fact that we are not considering the effect of the liquid and the container

on the signal in isolation, but are looking at the entire sensing domain as a whole. We first obtain the distribution of the complex permittivity in the sensing domain, and then extract the liquid information from the imaging result. Even though the container may have an effect on estimating the complex permittivity of the container edge region, it does not affect the overall complex permittivity profile of the liquid, which allows it to maintain good performance when the material of the container is different.

10.2.3 Impact of container shape. In order to test the influence of different shapes of cups on liquid identification, we select 4 containers of different shapes for testing, including the base is circular and the ground is rectangular. Their volumes vary from 350 ml to 500 ml. In addition, we noticed that due to the change of the complex permittivity distribution, the impact on the signal will change when the container with a rectangular bottom surface is rotated by a certain angle. Therefore, we regard the rectangular container itself and its container rotated $\frac{\pi}{4}$ as two kinds of containers during the experiment. For each container, we selected data from the test dataset with 10 random positions for evaluation. The data for each position contains 8 liquids. Having obtained the liquid features, we can adopt a simple K-Nearest Neighbors algorithm (K=1) to identify those liquids.

The identification results are shown in Fig. 14. It achieves an accuracy over 89%, which shows that *LiquImager* can identify liquids with high accuracy despite the shape of the containers. In the process of the signal from the transmitting antenna to the receiving antenna, a beam of electromagnetic waves is actually transmitted in space instead of a ray, which complicates the effect of centimeter-scale containers on the signal. In order to describe the influence of the medium on the electromagnetic signal as accurately as possible, we construct the electric field scattering model based on Maxwell's equations, which makes it have a finer sensing ability for small containers. Furthermore, we note that containers with rectangular bottoms show higher accuracy than cylindrical containers. We believe this is due to the former being larger at 500 ml than that of the latter (350 ml). The larger base area makes the scattering effect of the solution on the signal more significant, which facilitates fine-grained material identification.

10.2.4 Impact of container position. In actual deployment, the position of the liquid to be tested is often not fixed. We therefore test *LiquImager*'s ability to identify liquids placed in different locations. We randomly selected 10 positions from the data in the test set, which are shown in Fig. 15a. The data for each position contains 4 different containers and 8 different liquids. Having obtained the liquid features, we adopt a K-Nearest Neighbors classifier (K=1) to identify them.

The results are shown in Fig. 15b. The position represented by the box is indicated by color. For example, the point represented by *Pos1* is the upper right point in Fig. 15a. It achieves an accuracy over 89%, which shows that *LiquImager* can identify liquids with high accuracy despite the position of the containers. However, we noticed that the accuracy of liquid identification is low at locations far from the receiving antenna (such as Pos 4 and Pos 9). We believe this is due to the fact that when the target is farther away from the receiving antenna, the scattering strength of the signal becomes weaker, which makes it difficult for us to estimate the scattering data using the received CSI (Sec. 5.1).

10.3 Container Imaging

Next, we evaluate *LiquImager*'s imaging performance on four common containers. We expect it to display both the shape of the container and the material of the liquid in one image. Note that the output of the image enhancement network (*LiqU-Net*, introduced in Sec. 6) we designed is a grayscale image, and different grayscale values represent different liquids. In order to improve the recognition, we defined 8 colors to represent 8 different liquids, which are shown in Table 1.

We choose from 4 different containers and 8 different liquids. Each case contains data for 49 positions. We then perform edge detection and color transformation on the *LiqU-Net* output image. Specifically, we first perform
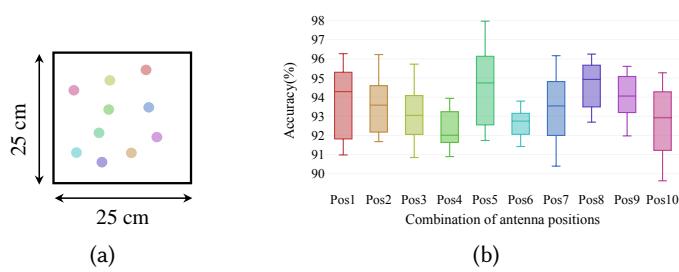
(a)

(b)

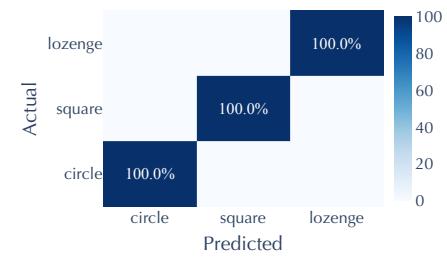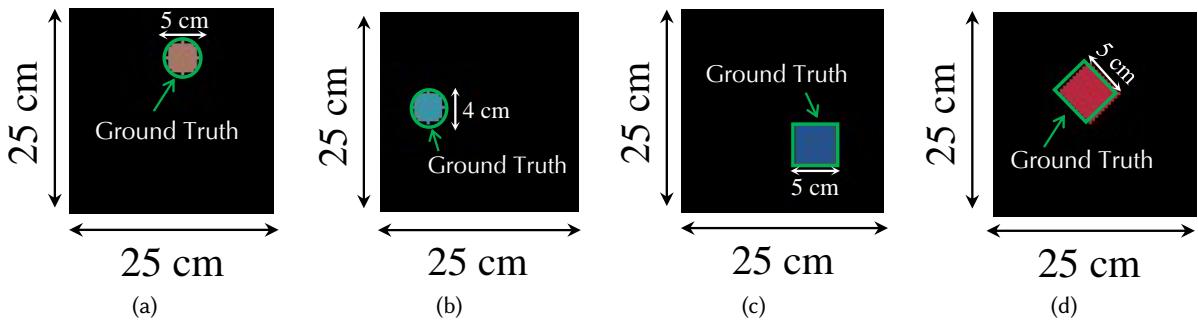Fig. 15. Performance when the liquids are placed at different positions.

Fig. 16. Performance for identifing shape.

(a)　　　　　　　(b)　　　　　　　(c)　　　　　　　(d)

Fig. 17. *LiquImager* imaging results for (a) oil in a circular container, (b) cola in a circular container, (c) water in a square container, and (d) sprite in a lozenge container.

edge detection, and then sort the edge coordinates after normalization. We use the sorted coordinates as features for shape classification. Then, after identifying the material according to the method described in Sec. 7, we replace the pixels in the material area for easy distinction.

The final imaging result is shown in Fig. 17. For containers of different widths and shapes, *LiquImager* can accurately image them. We believe that this is mainly due to the fact that the *LiqU-Net* network we built can better enhance the image quality. In addition, in order to meet the needs of material identification, different liquids present different colors, which correspond to the colors defined in Table 1. Edge detection is a relatively easy task after mapping different materials (liquid and air) into different pixels according to the distribution of complex permittivity. As shown in Fig. 16, *LiquImager* can identify the shape of a container with 100% accuracy. This means that despite the rotation of the container with a rectangular bottom, we can still image it well.

10.4 Boundary of Liquid Identification Capability

We design a series of experiment to test the boundaries of the *LiquImager*'s ability for liquid identification. Specifically, we test *LiquImager*'s ability to identify oil-based liquids, the effect of liquid height on liquid identification, and its ability to identify liquids held in extremely small containers (such as 20 ml test tubes).

(1) *We test the* LiquImager*'s ability to identify liquids that are nearly lossless media.* We select four cooking oil, including soybean oil, canola oil, peanut oil, and corn oil for testing. Both of them are all purchased from supermarkets. The collection of data is similar to that introduced in Sec. 10.1 and we use 4 types of containers,
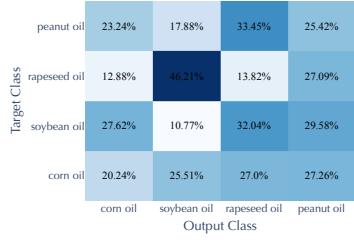
Fig. 18. *LiquImager* is unable to differentiate between the 4 vegetable oils
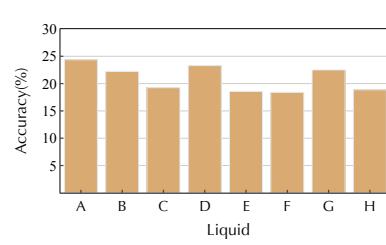

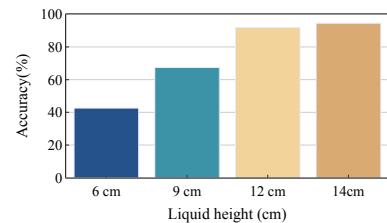Fig. 19. Identification with tube.

Fig. 20. Identification performance for liquid height.

which are shown in the Fig. 11. To facilitate data collection, we randomly selected 5 positions in the area to place containers. We partition the dataset using five-fold cross-validation and use the KNN classifier for liquid identification.

The confusion matrix is shown in Fig. 18. We find that *LiquImager*'s ability to identify lossless media is weak. We believe that the reason stems from two aspects. Specifically, on the one hand, their effect on the signal is close to that of air. The relative complex permittivity of these oils is about 2 [5], which is very close to the relative complex permittivity of air of 1. In comparison, the relative complex permittivity of water is about $73 + 18j$. Note that commercial Wi-Fi' devices have very strong noise, and weak scattered signals are easily drowned out. On the other hand, their relative complex permittivity is very close. For example, the difference between the relative complex permittivity of peanut oil and soybean oil is only 0.03 [5], which makes the difference in their influence on Wi-Fi signals so small that it is difficult to distinguish them. Similarly, it performed poorly at identifying different brands of pure milk. Specifically, we chose three types of pure milk, which come from (M1) Mengniu, (M2) Yili, and (M3) Deluxe. Experimental results show that LiquImager's identification accuracy for different types of pure milk is only 45%.

(2) LiquImager *performs poorly at identifying liquids held in extremely small containers due to the weak scatter signal.* As shown in Sec. 10.2, it can perform liquid identification with an accuracy of over 90% for containers over 350 ml. Although the volume of most containers in daily scenarios exceeds 350 ml, there are still some containers with small volumes, such as test tubes, which are usually less than 20 ml. We test the performance of *LiquImager* on very small containers. In particular, we place a 10 ml test tube on the rack and fill it with liquid for identification. We test 8 different liquids using a 5-fold cross-validation approach. The result is shown in the Fig. 19. We find that *LiquImager* is less able to discriminate between different liquids. We believe the reason is that the volume of the medium is so small that the resulting scattered signal is weak. For the situation where the signal-to-noise ratio (SNR) of the received signal is very poor, even if we use a deep learning network to enhance the signal, it cannot achieve high accuracy.

(3) *In order to effectively identify the liquid, the liquid height needs to exceed* 12 cm. To test the effect of liquids on the accuracy of liquid identification, we select the glass containers (350 ml) with different liquid heights, including 6 cm, 9 cm, 12 cm, and 14 cm. We select 8 liquids for identification. For the convenience of data collection, we select 5 random positions for each liquid. Having obtained the liquid features, we can adopt a K-Nearest Neighbors algorithm to differentiate those liquids. We use five-fold cross-validation for testing. The results are shown in Fig. 20. We find that in order to ensure high recognition accuracy, the height of liquid should be greater than 12 cm. We believe the reason is that the actual electric field scattering deviates from our model when the liquid height is small. Specifically, when we construct the electric field scattering model, we only consider the electromagnetic wave whose direction is perpendicular to the bottom surface of the container, which is a good approximation when the height of the medium is large. When the height of the medium is small, its two

bottom surfaces may refract the incident wave so as to generate electric fields in other directions, which leads to inaccurate identification results. However, we note that in actual scenes there are many liquids with a height higher than 12 cm. Besides, many solid targets (such as books, etc.) also meet this condition, which makes the system still have many potential applications. As a result, *LiquImager* still has a chance to work in many scenes. In future work, we will actively explore more general models for less restrictive sensing.

(4) LiquImager *cannot identify liquids in metal containers.* We choose a thermos cup made of 304 stainless steel and aluminum tea box as the container. Their diameter is approximately 6 cm. We selected 8 liquids as liquids to be tested, including water, cola, alcohol, lemon tea drink, oil, pepsi, sprite, and 1% alcohol solution. For the same liquid, we independently collected it 30 times, and then used the features output by the system for material identification. We find it almost difficult for the system to identify different liquids. The average identification accuracy is less than 20%. We believe this is because the metal container can be approximated as a Faraday cage [2], which makes it difficult for electromagnetic waves to enter the interior of the container. As a result, it's difficult to utilize Wi-Fi signals to sense the situation inside the container. We add relevant instructions in the revised version to more clearly describe the boundaries of system capabilities.

11 RELATED WORK

11.1 Wi-Fi based sensing

In recent years, the shipment of Wi-Fi devices has exceeded 30B, and 13B devices are already in user, which provide ubiquitous Wi-Fi coverage. Benefiting from ubiquitous coverage and low cost, researchers are actively exploring the application of Wi-Fi in indoor sensing, including health monitoring [26, 54], person identification [22, 40], occupancy estimation [16, 23], among other applications. Most relevant to our work are material identification and object imaging.

11.2 Material Identification

11.2.1 Traditional liquid identification methods. Traditionally, material identification requires the specialized equipment to provide data [35, 37], which makes them often costly. The basic principle of these methods is that the dielectric properties change with the type of material, which will have different effects on the incident wave. In addition to expensive equipment, these scheme usually require the immersion of a probe in a liquid to collect the signal, which will contaminate the liquid. These restrictions make it unsuitable for ubiquitous applications in daily life.

11.2.2 RF-based liquid identification method. In the recent years, lots of excellent inexpensive solutions based on wireless devices have been proposed, which facilitate the ubiquitous application of material identification.

As one of the most widely used frequency bands, there are numerous *centimeter wave-based* devices in real life, including Wi-Fi, RFID, UWB, and *etc. Based on these devices, many excellent works have been proposed for perception.* With an RFID tag that stuck on the surface of the material, Tagtag [46] construct a high-precision material identification system. Due to the existence of near-field coupling, the impedance of the RFID tag will vary with the liquid, which is used to construct the liquid signature. Similarly, RF-EATS [9] tries to identify the effect of materials on the backscatter tag's impedance for material recognition. However, these solutions often require placing tags on the target, which brings additional deployment costs. Tagscan [39] and LiquID [5] try a non-contact solution for liquid sensing. Specifically, they placed the liquid to be tested in the RF link, and extracted the unique characteristics of the liquid by analyzing the effect of the liquid on the received signal. They achieve a identification accuracy of more than 90% on 10 and 33 liquids to be tested, respectively. However, they require placing the liquid in a specific container, which limits their ubiquitous application. LiqRay [33]

---

[2]https://en.wikipedia.org/wiki/Faraday_cage

Table 2. Comparison of Liquid Identification Methods with Wireless Signals.

Method	Wavelength of the wireless signal	Deployment based on COTS devices	No specific container required	Imaging liquids	Container
TagTag [46]	centimeter level	yes	no[1]	no	5500 ml
RF-EATS [9]	centimeter level	yes	no	no	N/A
TagScan [39]	centimeter level	yes	no	yes	9000 ml
WiMi[6]	centimeter level	yes	yes	yes	3700 ml
LiquID [5]	centimeter level	yes	no	no	2700 ml
LiqRay [33]	centimeter level	no	yes	no	7200 ml
PackquID [34]	centimeter level	no	yes	no	350 ml
FG-LiquID [25]	millimeter level	yes	no	no	400 ml
mSense [44]	millimeter level	yes	yes	no	N/A
LiquImager	centimeter level	yes	yes	yes	350 ml

[1] no means not mentioned in the paper.
[2] N/A means the method is not applicable.

and PackquID [34] try to identify the liquid with different container shape, but it is unclear how well they will perform on widely used commercial devices. However, a significant number of container dimensions are in proximity to or even smaller than the Wi-Fi wavelength. These works often do not involve the perception of centimeter-level media, which is very common in daily life.

Due to its fine-grained sensing capability, millimeter-wave radar-based technology has received more and more attention. In terms of material identification, mSense [44] utilize millimeter wave radar to achieve high-precision (90% accuracy) identification of 5 materials in mobile scenes. FG-LiquID [25] designs a novel neural network for sensing liquids using millimeter-wave radar, which can identify 30 different liquids in a fine-grained manner. But they haven't explored imaging. If we could image the container while identifying the liquid, it would open up even more possibilities for ubiquitous sensing.

11.2.3 Optical and camera-based liquid detection method. Recent research on optical and camera-based liquid detection method has many different problem-solving theories and meaningful applications [14, 27, 52]. However, limited by the particularity of liquid identification, the application of the camera-based solution is limited. On the one hand, a large number of containers are opaque, and on the other hand, many liquids have the same color, such as water and alcohol. Both of them bring great difficulties to camera-based material identification schemes.

11.3 Object Imaging.

Recently, there are a amount of excellent work being devoted for imaging objects based on COTS Wi-Fi devices. In general, much work relies on the standard back-propagation imaging method [13, 30, 38, 53, 57], which utilizes measurements on a Rx grid to form an image of an object by tracing the received signal back to different locations in space, but the resolution is lower. This makes them difficult to apply to the imaging of centimeter-scale containers

12 CONCLUSION AND FUTURE WORK

Conclusion. In this paper, we propose *LiquImager*, which can image centimeter-scale containers and perform container shape-independent and position-independent liquid identification. Due to diffraction phenomena, the effect of centimeter-scale containers on the signal cannot be approximated by conventional ray tracing models. Starting directly from Maxwell's equations, we construct an electric field scattering model to describe the effect

of centimeter-scale containers on the signal. For material identification and object imaging, we do not consider them individually. Instead, based on our model, we attempt to estimate the complex permittivity distribution of the sensing region using numerous measurements. In order to suppress noise, we propose *LiqU-Net* for image enhancement. We deployed our system on COTS Wi-Fi devices. For a centimeter-scale container randomly placed in an area of 25 cm × 25 cm, it can accurately identify the liquid in it with more than 90%. In terms of container imaging, *LiquImager* can accurately determine the edge of the container for 4 types of containers with a volume less than 500 ml. We believe *LiquImager* can bring more possibilities for ubiquitous sensing.

Future work. We hope to use more accurate sensing equipment to obtain fine-grained scattering information to complete the identification of the lossless medium. Since *LiquImager* identifies liquids based on the scattering field, it is difficult to distinguish between non-consumable media (*e.g* vegetable oils). However, previous work has shown that the sensing accuracy of many wireless devices is better than that of Wi-Fi devices, which provides us with the possibility of accurate signal acquisition.

LiquImager uses the 5300 NIC for data acquisition, but it has been discontinued. Since *LiquImager* operates efficiently using only CSI data from the NIC and does not use the unique features of the NIC. In addition, the Wi-Fi 5 protocol used by the 5300 network card is compatible with most new network cards. Therefore, in principle, this approach shows full feasibility for compatibility with newer devices. In addition, we note that advanced CSI acquisition tools (such as PicoScense[3]) are compatible with some new network cards (such as Intel AX210). Moreover, some notable researches [43] have delved into the potential of estimating CSI data from readily accessible sources of information, such as BFI. We will actively try to do deploy evaluations on these modern network cards in the feature work.

We note that WiFi devices are still the most widely deployed wireless devices. Many shopping malls, rooms, and even mobile phones have WiFi devices. If we can complete some sensing tasks based on WiFi signals, its audience will be very wide. But as the reviewer pointed out, sensing based on WiFi devices still needs to be improved in terms of capabilities and robustness. In future work, in addition to exploring WiFi sensing theories and models, we will also actively try to use more professional signals such as millimeter waves, ultrasound, visible light, and even quantum for fine-grained sensing.

ACKNOWLEDGMENTS

REFERENCES

[1] Md Zahangir Alom, Chris Yakopcic, Mahmudul Hasan, Tarek M. Taha, and Vijayan K. Asari. 2019. Recurrent residual U-Net for medical image segmentation. *Journal of Medical Imaging* 6, 1 (March 2019), 014006. https://doi.org/10.1117/1.JMI.6.1.014006 Publisher: SPIE.

[2] Xudong Chen. 2018. *Computational methods for electromagnetic inverse scattering.* John Wiley & Sons.

[3] Guoxuan Chi, Zheng Yang, Jingao Xu, Chenshu Wu, Jialin Zhang, Jianzhe Liang, and Yunhao Liu. 2022. Wi-drone: wi-fi-based 6-DoF tracking for indoor drone flight control. In *Proceedings of the 20th annual international conference on mobile systems, applications and services.* 56–68.

[4] Tie Jun Cui, Weng Cho Chew, Xiao Xing Yin, and Wei Hong. 2004. Study of resolution and super resolution in electromagnetic imaging for half-space problems. *IEEE Transactions on Antennas and Propagation* 52, 6 (2004), 1398–1411. https://doi.org/10.1109/TAP.2004.829847

[5] Ashutosh Dhekne, Mahanth Gowda, Yixuan Zhao, Haitham Hassanieh, and Romit Roy Choudhury. 2018. Liquid: A wireless liquid identifier. In *Proceedings of the 16th annual international conference on mobile systems, applications, and services.* 442–454.

[6] Chao Feng, Jie Xiong, Liqiong Chang, Ju Wang, Xiaojiang Chen, Dingyi Fang, and Zhanyong Tang. 2019. WiMi: Target Material Identification with Commodity Wi-Fi Devices. IEEE. https://doi.org/10.1109/icdcs.2019.00075

---

[3]https://ps.zpj.io/

[7] Ruiyang Gao, Wenwei Li, Yaxiong Xie, Enze Yi, Leye Wang, Dan Wu, and Daqing Zhang. 2022. Towards Robust Gesture Recognition by Characterizing the Sensing Quality of WiFi Signals. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 6, 1 (2022), 1–26.

[8] David J. Griffiths. 2005. Introduction to electrodynamics. *American Journal of Physics* 73, 6 (2005), 574–574. https://doi.org/10.1119/1.4766311 tex.eprint: https://doi.org/10.1119/1.4766311.

[9] Unsoo Ha, Junshan Leng, Alaa Khaddaj, and Fadel Adib. 2020. Food and liquid sensing in practical environments using rfids. In *Proceedings of the 17th USENIX Symposium on Networked Systems Design and Implementation (NSDI'20)*.

[10] Feiyu Han, Chengchen Wan, Panlong Yang, Hao Zhang, Yubo Yan, and Xiang Cui. 2020. ACE: Accurate and automatic CSI error calibration for wireless localization system. In *2020 6th International Conference on Big Data Computing and Communications (BIGCOM)*. IEEE, 15–23.

[11] Richard D Haracz, Leonard D Cohen, and Ariel Cohen. 1985. Scattering of linearly polarized light from randomly oriented cylinders and spheroids. *Journal of Applied Physics* 58, 9 (1985), 3322–3327.

[12] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. 2016. Deep Residual Learning for Image Recognition. 770–778. https://openaccess.thecvf.com/content_cvpr_2016/html/He_Deep_Residual_Learning_CVPR_2016_paper.html

[13] Donny Huang, Rajalakshmi Nandakumar, and Shyamnath Gollakota. 2014. Feasibility and limits of wi-fi imaging. In *Proceedings of the 12th ACM conference on embedded network sensor systems*. 266–279.

[14] Qianyi Huang, Zhice Yang, and Qian Zhang. 2018. Smart-U: smart utensils know what you eat. In *IEEE INFOCOM 2018-IEEE Conference on Computer Communications*. IEEE, 1439–1447.

[15] Zhongling Huang, Mihai Datcu, Zongxu Pan, and Bin Lei. 2020. Deep SAR-Net: Learning Objects from Signals. *ISPRS Journal of Photogrammetry and Remote Sensing* 161 (March 2020), 179–193. https://doi.org/10.1016/j.isprsjprs.2020.01.016

[16] Osama Talaat Ibrahim, Walid Gomaa, and Moustafa Youssef. 2019. CrossCount: A deep learning system for device-free human counting using WiFi. *IEEE Sensors Journal* 19, 21 (2019), 9921–9928. Publisher: IEEE.

[17] Fabian Isensee, Jens Petersen, Andre Klein, David Zimmerer, Paul F. Jaeger, Simon Kohl, Jakob Wasserthal, Gregor Koehler, Tobias Norajitra, Sebastian Wirkert, and Klaus H. Maier-Hein. 2018. nnU-Net: Self-adapting Framework for U-Net-Based Medical Image Segmentation. https://doi.org/10.48550/arXiv.1809.10486 arXiv:1809.10486 [cs].

[18] J. D. Jackson and Ronald F. Fox. 1999. Classical electrodynamics, 3rd ed. *American Journal of Physics* 67, 9 (Sept. 1999), 841–842. https://doi.org/10.1119/1.19136 tex.eprint: https://pubs.aip.org/aapt/ajp/article-pdf/67/9/841/7528211/841_2_online.pdf.

[19] Zhiping Jiang, Tom H. Luan, Xincheng Ren, Dongtao Lv, Han Hao, Jing Wang, Kun Zhao, Wei Xi, Yueshen Xu, and Rui Li. 2022. Eliminating the Barriers: Demystifying Wi-Fi Baseband Design and Introducing the PicoScenes Wi-Fi Sensing Platform. *IEEE Internet of Things Journal* 9, 6 (2022), 4476–4496. https://doi.org/10.1109/JIOT.2021.3104666

[20] Jian-Ming Jin. 2015. *Theory and Computation of Electromagnetic Fields*. John Wiley & Sons. Google-Books-ID: Q0UxBwAAQBAJ.

[21] Kyong Hwan Jin, Michael T. McCann, Emmanuel Froustey, and Michael Unser. 2017. Deep Convolutional Neural Network for Inverse Problems in Imaging. *IEEE Transactions on Image Processing* 26, 9 (Sept. 2017), 4509–4522. https://doi.org/10.1109/TIP.2017.2713099 Conference Name: IEEE Transactions on Image Processing.

[22] Belal Korany, Chitra R Karanam, Hong Cai, and Yasamin Mostofi. 2019. XModal-ID: Using WiFi for through-wall person identification from candidate video footage. In *The 25th annual international conference on mobile computing and networking*. 1–15.

[23] Belal Korany and Yasamin Mostofi. 2021. Counting a stationary crowd using off-the-shelf wifi. In *Proceedings of the 19th annual international conference on mobile systems, applications, and services*. 202–214.

[24] John Daniel Kraus and Ronald J. Marhefka. 2002. *Antennas for all applications*. https://ui.adsabs.harvard.edu/abs/2002aaa..book.....K Publication Title: Antennas for all applications ADS Bibcode: 2002aaa..book.....K.

[25] 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. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 5, 3 (Sept. 2021), 1–27. https://doi.org/10.1145/3478075

[26] Xuefeng Liu, Jiannong Cao, Shaojie Tang, Jiaqi Wen, and Peng Guo. 2015. Contactless respiration monitoring via off-the-shelf WiFi devices. *IEEE Transactions on Mobile Computing* 15, 10 (2015), 2466–2479. Publisher: IEEE.

[27] Hidenori Matsui, Takahiro Hashizume, and Koji Yatani. 2018. Al-light: An alcohol-sensing smart ice cube. *Proceedings of the ACM on interactive, mobile, wearable and ubiquitous technologies* 2, 3 (2018), 1–20.

[28] Sophocles J Orfanidis. 2002. *Electromagnetic waves and antennas*. Rutgers University New Brunswick, NJ.

[29] Anurag Pallaprolu, Belal Korany, and Yasamin Mostofi. 2022. Wiffract: A New Foundation for RF Imaging via Edge Tracing. In *Proceedings of the 28th Annual International Conference on Mobile Computing And Networking*. ACM, Sydney NSW Australia, 255–267. https://doi.org/10.1145/3495243.3514261

[30] Anurag Pallaprolu, Belal Korany, and Yasamin Mostofi. 2022. Wiffract: a new foundation for RF imaging via edge tracing. In *Proceedings of the 28th annual international conference on mobile computing and networking*. 255–267.

[31] Andrew F Peterson, Scott L Ray, Raj Mittra, Institute of Electrical, and Electronics Engineers. 1998. *Computational methods for electromagnetics*. Vol. 351. IEEE press New York.

[32] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. 2015. U-Net: Convolutional Networks for Biomedical Image Segmentation. In *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015 (Lecture Notes in Computer Science)*, Nassir Navab, Joachim Hornegger, William M. Wells, and Alejandro F. Frangi (Eds.). Springer International Publishing, Cham, 234–241. https://doi.org/10.1007/978-3-319-24574-4_28

[33] Fei Shang, Panlong Yang, Yubo Yan, and Xiang-Yang Li. 2022. LiqRay: non-invasive and fine-grained liquid recognition system. In *Proceedings of the 28th annual international conference on mobile computing and networking*. 296–309.

[34] Fei Shang, Panlong Yang, Yubo Yan, and Xiang-Yang Li. 2023. PackquID: In-packet liquid identification using RF signals. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 6, 4 (2023), 1–27. Publisher: ACM New York, NY, USA.

[35] Monika Szymańska-Chargot, Justyna Cybulska, and Artur Zdunek. 2011. Sensing the structural differences in cellulose from apple and bacterial cell wall materials by Raman and FT-IR spectroscopy. *Sensors* 11, 6 (2011), 5543–5560. Publisher: Molecular Diversity Preservation International (MDPI).

[36] John Terry and Juha Heiskala. 2002. *OFDM Wireless LANs: A Theoretical and Practical Guide*. Sams Publishing. Google-Books-ID: p_rMm2l237gC.

[37] Georgios Tsiminis, Fenghong Chu, Stephen C Warren-Smith, Nigel A Spooner, and Tanya M Monro. 2013. Identification and quantification of explosives in nanolitre solution volumes by Raman spectroscopy in suspended core optical fibers. *Sensors* 13, 10 (2013), 13163–13177. Publisher: Molecular Diversity Preservation International (MDPI).

[38] Chen Wang, Jian Liu, Yingying Chen, Hongbo Liu, and Yan Wang. 2018. Towards in-baggage suspicious object detection using commodity wifi. In *2018 IEEE Conference on Communications and Network Security (CNS)*. IEEE, 1–9.

[39] Ju Wang, Jie Xiong, Xiaojiang Chen, Hongbo Jiang, Rajesh Krishna Balan, and Dingyi Fang. 2017. TagScan: Simultaneous target imaging and material identification with commodity RFID devices. In *Proceedings of the 23rd annual international conference on mobile computing and networking*. 288–300.

[40] Wei Wang, Alex X Liu, and Muhammad Shahzad. 2016. Gait recognition using wifi signals. In *Proceedings of the 2016 ACM international joint conference on pervasive and ubiquitous computing*. 363–373.

[41] Xuanzhi Wang, Kai Niu, Jie Xiong, Bochong Qian, Zhiyun Yao, Tairong Lou, and Daqing Zhang. 2022. Placement matters: Understanding the effects of device placement for WiFi sensing. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 6, 1 (2022), 1–25.

[42] Zhun Wei and Xudong Chen. 2019. Deep-Learning Schemes for Full-Wave Nonlinear Inverse Scattering Problems. *IEEE Transactions on Geoscience and Remote Sensing* 57, 4 (April 2019), 1849–1860. https://doi.org/10.1109/TGRS.2018.2869221 Conference Name: IEEE Transactions on Geoscience and Remote Sensing.

[43] Chenhao Wu, Xuan Huang, Jun Huang, and Guoliang Xing. 2023. Enabling Ubiquitous WiFi Sensing with Beamforming Reports. In *Proceedings of the ACM SIGCOMM 2023 Conference*. 20–32.

[44] Chenshu Wu, Feng Zhang, Beibei Wang, and K. J. Ray Liu. 2020. mSense: Towards Mobile Material Sensing with a Single Millimeter-Wave Radio. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 4, 3 (2020), 106:1–106:20. https://doi.org/10.1145/3411822

[45] Li-Ye Xiao, Jiawen Li, Feng Han, Wei Shao, and Qing Huo Liu. 2020. Dual-Module NMM-IEM Machine Learning for Fast Electromagnetic Inversion of Inhomogeneous Scatterers With High Contrasts and Large Electrical Dimensions. *IEEE Transactions on Antennas and Propagation* 68, 8 (Aug. 2020), 6245–6255. https://doi.org/10.1109/TAP.2020.2990222 Conference Name: IEEE Transactions on Antennas and Propagation.

[46] Binbin Xie, Jie Xiong, Xiaojiang Chen, Eugene Chai, Liyao Li, Zhanyong Tang, and Dingyi Fang. 2019. Tagtag: Material sensing with commodity RFID. In *Proceedings of the 17th Conference on Embedded Networked Sensor Systems (SenSys '19)*. Association for Computing Machinery, New York, NY, USA, 338–350. https://doi.org/10.1145/3356250.3360027

[47] Yaxiong Xie, Zhenjiang Li, and Mo Li. 2015. Precise power delay profiling with commodity WiFi. In *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking*. 53–64.

[48] Jie Xiong and Kyle Jamieson. 2013. Arraytrack: A fine-grained indoor location system. tex.organization: Usenix.

[49] He Ming Yao, Lijun Jiang, and Wei E. I. Sha. 2020. Enhanced Deep Learning Approach Based on the Deep Convolutional Encoder–Decoder Architecture for Electromagnetic Inverse Scattering Problems. *IEEE Antennas and Wireless Propagation Letters* 19, 7 (July 2020), 1211–1215. https://doi.org/10.1109/LAWP.2020.2995455 Conference Name: IEEE Antennas and Wireless Propagation Letters.

[50] 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 (UIST '16)*. Association for Computing Machinery, New York, NY, USA, 833–841. https://doi.org/10.1145/2984511.2984515

[51] Fan Yin, Chang Chen, Weidong Chen, Wen Qiao, and Jian Guan. 2020. Superresolution quantitative imaging based on superoscillatory field. *Opt. Express* 28, 5 (Mar 2020), 7707–7720. https://doi.org/10.1364/OE.384866

[52] Shichao Yue and Dina Katabi. 2019. Liquid testing with your smartphone. In *Proceedings of the 17th Annual International Conference on Mobile Systems, Applications, and Services*. 275–286.

[53] Shengkai Zhang, Gaoyang Liu, S Wan, et al. 2021. WiImage: Fine-grained Human Imaging from Commodity WiFi. (2021).

[54] Mingmin Zhao, Shichao Yue, Dina Katabi, Tommi S Jaakkola, and Matt T Bianchi. 2017. Learning sleep stages from radio signals: A conditional adversarial architecture. In *International conference on machine learning*. 4100–4109. tex.organization: PMLR.

[55] Zongwei Zhou, Md Mahfuzur Rahman Siddiquee, Nima Tajbakhsh, and Jianming Liang. 2018. UNet++: A Nested U-Net Architecture for Medical Image Segmentation. In *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support (Lecture Notes in Computer Science)*, Danail Stoyanov, Zeike Taylor, Gustavo Carneiro, Tanveer Syeda-Mahmood, Anne Martel, Lena Maier-Hein, João Manuel R.S. Tavares, Andrew Bradley, João Paulo Papa, Vasileios Belagiannis, Jacinto C. Nascimento, Zhi Lu, Sailesh Conjeti, Mehdi Moradi, Hayit Greenspan, and Anant Madabhushi (Eds.). Springer International Publishing, Cham, 3–11. https://doi.org/10.1007/978-3-030-00889-5_1

[56] Hongzi Zhu, Yiwei Zhuo, Qinghao Liu, and Shan Chang. 2018. π-splicer: Perceiving accurate CSI phases with commodity WiFi devices. *IEEE Transactions on Mobile Computing* 17, 9 (2018), 2155–2165.

[57] Yang Zi, Wei Xi, Li Zhu, Fan Yu, Kun Zhao, and Zhi Wang. 2019. Wi-Fi imaging based segmentation and recognition of continuous activity. In *Collaborative Computing: Networking, Applications and Worksharing: 15th EAI International Conference, CollaborateCom 2019, London, UK, August 19-22, 2019, Proceedings 15*. Springer, 623–641.