Environment-independent In-baggage Object Identification Using WiFi Signals

Cong Shi*, Tianming Zhao†, Yucheng Xie‡, Tianfang Zhang*, Yan Wang†, Xiaonan Guo‡, Yingying Chen*
*Rutgers University, New Brunswick, NJ 08901, USA
†Temple University, Philadelphia, PA 19122, USA
‡Indiana University-Purdue University Indianapolis, IN 46202, USA
Email: *{cs1421, tz203, yingche}@scarletmail.rutgers.edu, †{tum94362, y.wang}@temple.edu, ‡{yx11, xg6}@iupui.edu

***Abstract*—Low-cost in-baggage object identification is highly demanded in enhancing public safety and smart manufacturing. Existing approaches usually require specialized equipment and heavy deployment overhead, making them hard to scale for wide deployment. The recent WiFi-based approach is unsuitable for practical deployment as it did not address dynamic environmental impacts. In this work, we propose an environment-independent in-baggage object identification system by leveraging low-cost WiFi. We exploit the channel state information (CSI) to capture material and shape characteristics to facilitate fine-grained in-baggage object identification. A major challenge of building such a system is that CSI measurements are sensitive to real-world dynamics, such as different types of baggage, time-varying ambient noises and interferences, and different deployment environments. To tackle these problems, we develop WiFi features based on polarized directional antennas that can capture objects' material and shape characteristics. A convolutional neural network-based model is developed to constructively integrate the WiFi features and perform accurate in-baggage object identification. We also develop a material-based domain adaptation using adversarial learning to facilitate fast deployments in different environments. We conduct extensive experiments involving 14 representation objects, 4 types of bags in 3 different room environments. The results show that our system can achieve over 97% in the same environment, and our domain adaptation method can improve the object identification accuracy by 42% when the system is deployed in a new environment with little training.**

I. Introduction

In 2020, over 614 mass shootings happened in the U.S., resulting in $3,061$ victims, including 446 death and $2,515$ injuries [1]. Additionally, homemade bombs have led to 227 fatalities from 2012 to 2019 [2]. In most of these tragedies, the attackers carried dangerous objects (e.g., guns and homemade bombs) in bags or suitcases to public venues, such as hospitals, schools, theaters, parks, and launched the attacks without being noticed. Therefore, it is highly demanded to widely deploy in-baggage object identification systems in public places, which will allow first responders to detect concealed dangerous objects in public places and enhance public safety. In-baggage object identification can also facilitate smart inspection for product manufacture and delivery through verifying assortment and quantities in bags/packages without manually checking the enclosed products (e.g., gift bags, mailing packages, or any packed items). Such an automatic and non-intrusive quality inspection can significantly reduce labor costs and improve product manufacturing and delivery efficiency.

Existing methods for in-baggage object identification mainly rely on X-ray [3] and CT machinery [4]. While these methods are effective, they incur significant deployment overhead and financial costs. As such, only a few security-critical places, such as airports, courthouses, and police stations, are currently using these methods at their public entrances. Researchers have started exploring radio frequency (RF) signals to achieve material or object identification. The radio-frequency identification (RFID) [5], [6] technology has been proposed to determine the type of liquid in a container or measure an object's shape by attaching RFID tags to the target container or object. Ultra-wideband or mmWave signals have been exploited to determine an object's material [7] and shape [8] without attaching any sensors to the object. None of these methods applies to in-baggage object identification because they do not consider the impacts of bags.

Recent years have witnessed the emergence of using WiFi for various sensing applications, including generating images of human bodies or objects [9] and detecting materials [10]. Our study [11] has shown the initial success of in-baggage suspicious objects detection using WiFi. While this system can detect hidden suspicious objects made of metal or liquid, it is not fine-grained enough to differentiate hidden objects made of the same material. Such a capability is essential as it can significantly reduce false positives in practical scenarios where many everyday items in bags are metallic or liquid. Moreover, none of these works develops a general method that can be deployed in different environments (e.g., different room layouts and furniture placements) with low efforts.

This paper aims to develop an in-baggage object identification system that can facilitate wide deployment in various environments with low efforts. Toward this end, we utilize prevalent WiFi to build a low-cost non-intrusive sensing system that can recognize different objects in bags without using any additional sensors. Different from existing works, we extensively study the WiFi signals from directional antennas and develop unique features from WiFi measurements (i.e., channel state information (CSI)) to capture hidden objects' material and shape characteristics. Such fine-grained features allow our system to accurately identify different hidden objects even when they are made of the same material. In addition, we leverage deep learning technology to develop a generalized object classifier that can be easily adapted to various real-world

dynamics, including time-varying ambient noises, interference, and physical environmental changes (e.g., different furniture placement and room layouts). Our learning-based classifier models the relationships between objects made of the same material and significantly reduces the training effort when deploying the system in a new environment.

To realize such a system, we need to address many critical challenges. It is known that many factors in real environments can affect WiFi signals, for example, ambient noises and hardware distortion. Our system needs to identify effective features highly correlated with objects' material and shape characteristics and less sensitive to other impact factors. In addition, due to the open nature of wireless communications, the WiFi signals arriving at the receiver carry not only the object's characteristics but also various information of the surrounding environments, including the types of bags, WiFi interference, room layouts, and furniture placements. We refer to these factors as different domains and the signal pattern variations as domain shifts. Thus, our system should be robust to such domain shifts when it is deployed in different domains. Furthermore, to facilitate deployment on a large scale, our approach should have the capability to adapt to new deployment locations quickly and maintain a good performance with significantly reduced training efforts.

In this work, we propose exploiting polarized directional antennas to build a WiFi sensing platform that can enhance the WiFi signals for separating in-baggage objects' characteristics from the impacts of surrounding environments. Based on fine-grained CSI measurements, we derive the polarization feature and CSI complex difference feature for capturing objects' material and shape characteristics, respectively. A convolutional neural network (CNN) is developed to learn the abstract representations of objects and facilitate building an environment-independent in-baggage object identification model. A domain adaptation framework grounded on adversarial learning is designed to help transfer our object identification model to new domains and enable accurate object identification with a smaller dataset. Furthermore, we find that objects made of the same material usually share similar signal patterns compared to those made of different materials regardless of domains. Therefore, we design our domain adaption framework to leverage such inter-object relationships to further reduce the training effort for domain adaptation. We highlight our contributions in this work as follows:

- We demonstrate that WiFi signals provided by low-cost commodity devices can be exploited to accurately identify objects in baggage (e.g., laptops, clothes, bottled water) even for objects made of the same material.
- Our extensive studies on the WiFi signals from polarized directional antennas show that polarization characteristics and differences of CSI complex between two receiving antennas can effectively capture objects' material and shape characteristics. A CNN-based classifier is developed to extract object representations and facilitate environment-independent in-baggage object identification.

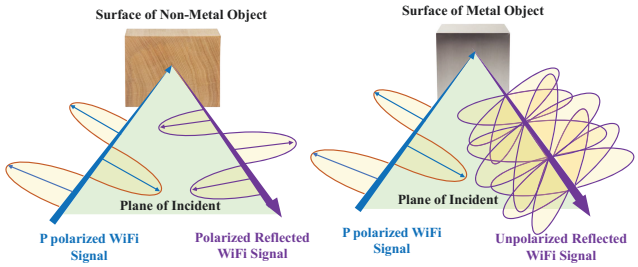


Fig. 1. Polarization effects of WiFi signals when reflecting from non-metal and metal surfaces.

- We exploit the inter-object relationships among objects of the same material to develop a material-based domain adaptation scheme. Our scheme mitigates the domain impacts from various real-world dynamics (i.e., different types of bags, time-varying interference, and physical environments). It allows adapting our trained in-baggage object identification model to a new domain with significantly reduced training efforts.
- We conduct extensive experiments involving 14 common in-baggage objects, four types of bags, at three different room environments. The results demonstrate the effectiveness of our system with over 97% object identification accuracy in the same environment. We also show that our domain adaptation method can improve the accuracy by 42% when the system is deployed in a new environment.

II. Related Work

Traditional methods on object identification can be categorized as vision-based and acoustic-based approaches. Vision-based methods [12], [13] use optical camera, hyper-spectral camera, and ToF camera to capture distinctive texture and geometry of the object for material and object identification. However, these vision-based solutions require a line-of-sight path between the camera and the object, and their performance can be degraded by surrounding lighting conditions. Besides leveraging the visual patterns, object sensing based on acoustic signals has been explored. For instance, AIM [14] utilizes the microphone and speaker on a smartphone to image objects in baggage. However, such a system requires moving a phone through a predefined trajectory to synthesize a virtual sensor array for imaging. Such complex manual operations make the system difficult to scale and prone to users' careless behaviors.

Recent studies have shown the potential of using various radio frequency (RF) signals to perform material identification or shape detection [8], [15]. For instance, LiquID [15] leverages ultra-wideband (UWB) signals propagating through a liquid container to measure the liquid's permittivity for liquid type identification. Zhu et al. [8] explore using 60GHz mmWave signals to determine the location, shape, and material of a target object through analyzing the strength of reflected signals. However, these systems rely on specialized equipment (e.g., antenna arrays and radar sensors), which adds extra cost and limits their large-scale deployment. Besides UWB and mmWave signals, RFID-based approaches have been explored for object sensing [5], [6]. For instance, TagScan [5] iden-

tifies the material type and detects the object shape through attaching RFID tags on the object's surface. However, such approaches still require specialized RFID transceivers to transmit/receive wireless signals, and their effectiveness on sensing in-baggage objects is not clear.

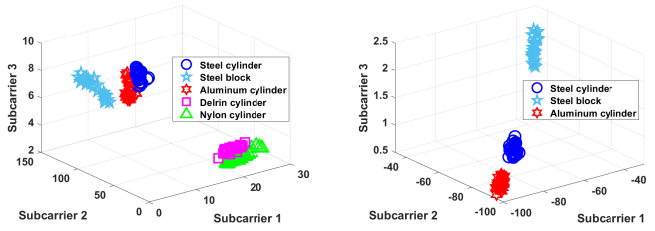
To remove the requirement on dedicated radar sensors, researchers explored the feasibility of using commodity WiFi for non-intrusive sensing, including activity recognition [16], vital sign monitoring [17], material identification and object shape detection [11], [18], [19]. For instance, E-eyes [16] is among the first studies to identify daily activity using home WiFi devices (e.g., desktops, refrigerators, access points). Liu *et al.* [17] first explore using wireless signals generated by commodity WiFi devices to perform vital sign monitoring (i.e., breathing and heart rate monitoring). Some recent studies also explored material and object shape detection with WiFi signals [11], [18]–[21]. Huang et al. [18] propose an object imaging system by using WiFi signals bouncing off of objects (e.g., leather couches and metallic shapes). WiMi [19] identifies liquid types based on the phase and amplitude variations of WiFi signals passing through the liquid. Our previous study [11] shows initial success in using WiFi signals to detect suspicious objects in baggage, such as liquid and metal objects. However, the designed system can not differentiate objects made of the same material (e.g., laptops and knives).

In this work, we design and develop a low-cost and portable system that leverages both material and shape characteristics in the WiFi signals to perform object identification. Different from existing studies, our system can differentiate even objects made of the same material. Furthermore, we address various challenges of real-world dynamics (i.e., adapting to physical environment changes) through developing a domain adaptation framework leveraging inter-object relationships among objects to reduce the training cost of the adaptation.

III. Preliminaries

A. WiFi Signal Polarization

It is known that WiFi signals propagate as the electromagnetic wave oscillating perpendicularly to the signal propagation direction. Polarized WiFi signals usually oscillate on a single plane (or two planes for dual-polarization), while the unpolarized WiFi signals could oscillate in all planes perpendicular to the propagation direction. The polarized WiFi signal can be decomposed into two orthogonal polarized directions, including p-polarization, with the oscillation direction on the plane of incident, and s-polarization, with the oscillation direction perpendicularly to the plane of incident. Upon bouncing off the object's surface, the polarization of WiFi signals can carry information reflecting the material of the object [22]. We illustrate the polarization effect in Figure 1. Assuming p-polarized WiFi signals hit the surface of a wooden object (i.e., with a non-metallic surface), the reflected signals are still polarized and contain both s- and p-polarization components. In contrast, a metallic object will depolarize the reflected WiFi signals, and the electromagnetic wave can vibrate at all directions perpendicular to the propagation direction. By measuring the polarization of the reflected WiFi signals, we can differentiate objects with different materials (e.g., a book made of fiber and a laptop made of metal). In this work, we utilize a pair of mutually perpendicular receiving antennas to capture the polarization effects of WiFi signals.



(a) Material characteristics (b) Shape characteristics

Fig. 2. Illustration of using CSI in WiFi signals to capture material and shape characteristics to differentiate objects.

To study the feasibility of using the polarization effects to differentiate materials, we collect CSI of WiFi signals reflected from 4 cylinders with the same shape but made of different materials (i.e., steel, aluminum, Delrin, and nylon). We also collect CSI of WiFi signals reflected from a steel block to study the impact of the shape. Based on the power ratio between CSI collected from two orthogonal receiving antennas, we derive the polarization feature to capture the polarization effects. The detailed method to extract the polarization features is elaborated in Section V-B. Figure 2 (a) shows the objects' polarization features extracted from 3 subcarriers. We observe that polarization features of metal objects (i.e., steel cylinder, steel block, and aluminum cylinder) are well separable from those of non-metal objects (i.e., Delrin and nylon cylinders). This result demonstrates the potential of using polarization features to differentiate metal and non-metal objects.

B. CSI Complex Difference

Although the WiFi signal polarization can be used to differentiate metal and non-metal objects, it is not sensitive to shape characteristics (e.g., Figure 2 (a) shows that the polarization features of the steel block and the steel cylinder are similar), making it difficult to differentiate objects made of the same material. To enable fine-grained object identification, we leverage the spatial information captured in a pair of receiving antennas and derive CSI complex difference, which characterizes the relative channel between two receiving antennas. The CSI complex difference at the i^{th} subcarrier can be formulated as:

$$\hat{H}_i = H_i^{t\leftrightarrow r_1}\left(H_i^{t\leftrightarrow r_2}\right)^* = \left|\hat{H}_i\right| e^{j\angle\hat{H}_i}, \tag{1}$$

where $H_i^{t\leftrightarrow r_1}$ and $H_i^{t\leftrightarrow r_2}$ denote the CSI from the links $t \leftrightarrow r_1$ and $t \leftrightarrow r_1$, respectively. $(.)^*$ denotes the complex conjugate. Equation 1 removes the shared internal hardware interference [23] and amplifies the minute channel distortions caused by the object's material and shape, such as attenuation and scattering. As shown in Figure 2 (b), the CSI complex differences derived from the WiFi signals reflected by different objects are more sensitive to the objects' shape. The distance between the steel cylinder and aluminum cylinder clusters is

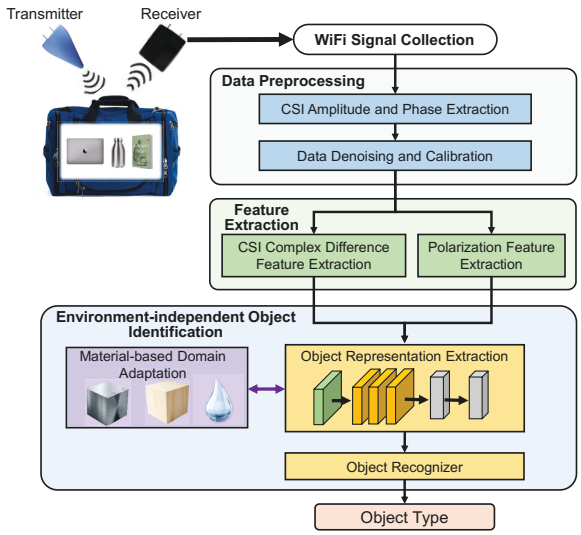
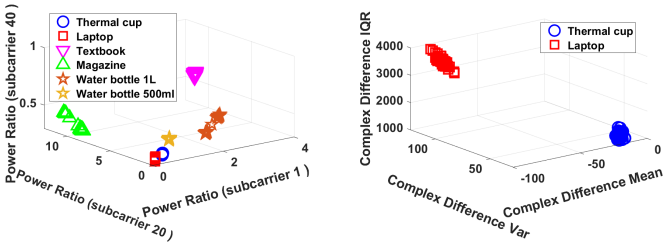


Fig. 3. System architecture.



(a) Polarization features (b) CSI complex difference

Fig. 4. Illustration of using polarization features and CSI complex difference statistics to capture material and shape characteristics of the object.

closer than the distance between the steel block and steel cylinder clusters, which indicates that the CSI complex difference encodes objects' shape characteristics.

IV. System Design

A. Challenges

Unreliable Material and Shape Characteristics. The WiFi signals reflected from target objects can be easily affected by channel condition changes and internal hardware noises, which cause unpredictable sampling time and frequency offsets in CSI measurements. To enable reliable in-baggage object identification, it is necessary to mitigate the impacts of such noises and reveal distinctive material and shape characteristics.

Real-world Dynamics. In real-world scenarios, the signals reflected from target objects can be impacted by various real-world dynamics, including different types of bags, noises and interferences from neighboring WiFi devices, and the physical environment changes. Such real-world dynamics can alter the multipath and change the signal patterns of CSI, resulting in false predictions. We need to mitigate the impacts of such dynamics to enable reliable object identification.

Significant Training Efforts. During deployment, it would be infeasible to train a machine learning model using all possible objects, which involves significant training efforts. It would be desirable to have a fast deployment method that only involves a minimal amount of data for training.

B. System Overview

To address the aforementioned challenges, we propose to extract material and shape characteristics in WiFi signals to identify in-baggage objects. Furthermore, to mitigate the impacts of various real-world dynamics (e.g., bags, ambient interference, and physical environments), we develop a domain adaptation framework based on adversarial training to reduce the training efforts. We illustrate the overview of our system in Figure 3. Particularly, our system utilizes a directional antenna to send horizontally polarized WiFi signals towards a target object. By using a pair of orthogonal antennas as the receiver, the amplitude, phase, and polarization of the received WiFi signals can capture rich material and shape characteristics. Our system then extracts CSI amplitude and phase from each OFDM subcarrier for data preprocessing. To remove the impacts of sampling time/frequency offsets introduced by the internal hardware of WiFi transmitter/receiver, which distorts the CSI phases, we calibrate the phase values of individual subcarrier by using a linear transformation method [23].

Next, our system extracts two sets of features from the calibrated CSI measurements, including polarization features and CSI complex difference statistics. The polarization features are the power ratios between two orthogonal receiving antennas, which characterizes the polarization direction changes induced by the material of the target object. To further differentiate objects with similar materials (e.g., a laptop and a thermal cup made of steel) and also capture shape characteristics, we derive CSI complex difference between two receiving antennas, which encodes spatial information of signal attenuation, scattering, reflection, and refraction. Based on the CSI complex difference, we extract a set of statistics features that capture signal patterns (e.g., signal attenuation, phase shifts) reflecting the object's material and shape.

Based on the extracted features, we develop a deep learning model to perform object identification. An object representation extractor based on a convolutional neural network (CNN) is used to learn object representations, which amplifies the distinctive material and shape characteristics. An object recognizer based on a fully connected neural network is then utilized to identify the object type. Furthermore, we develop a low-effort domain adaptation framework to mitigate various real-world dynamics. Our framework only requires the data of one object per material for the adaptation, which significantly reduces the training efforts during maintenance and deployment. Particularly, we find that the representation distances of different objects have similar patterns across different domains, and we define such distances as inter-object relationships. We leverage such relationships among objects of different materials and develop a domain discriminator to transfer the material characteristics as a whole from the source to the target domain. Furthermore, to ensure our system capable of differentiating objects of the same material in the target domain, we further transfer the shape characteristics learned in the source domain for the adaptation.

V. Feature Extraction

A. Data Denoising and Calibration

To extract reliable features for object identification, we first need to remove the impacts of hardware noises in CSI. Due to the imperfect synchronization, the CSI phase can be distorted by the sampling time offset (STO) and sampling frequency offset (SFO) between the transmitter and the receiver [23]. These two types of offsets introduce unpredictable time lag Δt and phase error $\Delta\theta$ to the phase of CSI, respectively. The distorted CSI phase can be formulated as:

$$\angle H_i = \widetilde{\angle H_i} + 2\pi f_i \Delta t + \Delta\theta, \tag{2}$$

where $\widetilde{\angle H_i}$ is the phase without distortions. i denotes the index of the subcarrier and f_i represents the frequency of the subcarrier. To calibrate the phase value, we first calculate the CSI magnitude and phase for all subcarriers. Then, we exploit a linear transformation-based method [23] to estimate these two types of errors:

$$\Delta t = \frac{\angle H_K - \angle H_1}{2\pi K}, \Delta\theta = \frac{1}{K}\sum_{i=1}^{K} \angle H_i, \tag{3}$$

where K is the total number of subcarriers (e.g., $K = 56$ for Atheros WiFi NICs). By subtracting $2\pi f_i \Delta t$ and $\Delta\theta$ from the phase of the i_{th} subcarrier, we can calibrate the phase values and mitigate the impacts of STO and SFO. Furthermore, due to the internal hardware imperfections (e.g., caused by the low-cost ADC in commodity WiFi devices), the magnitude and phase of CSI are very noisy. We apply a moving average filter upon the time-series magnitude and phase values to remove such noises. Given the magnitude and phase of denoised CSI, we reconstruct the complex form of CSI and use it to derive the CSI complex difference and polarization features.

B. Polarization Feature Extraction

As discussed in Section III, WiFi signals reflected from a target object encode the material characteristics in terms of signal polarization. Hence, we aim to first extract features that well capture such polarization characteristics. The polarization can be measured as the ratio of signal power between two orthogonal polarized antennas, which characterizes the oscillation direction of the electromagnetic wave while removing the amplitude attenuation due to signal propagation. We denote the CSI of the i^{th} subcarrier between a pair of transmitting antenna t and receiving antenna r as:

$$H_i^{t\leftrightarrow r} = \left|H_i^{t\leftrightarrow r}\right| e^{j\angle H_i^{t\leftrightarrow r}}, \tag{4}$$

where $|H_i^{t\leftrightarrow r}|$ and $\angle H_i^{t\leftrightarrow r}$ represent the amplitude and phase, respectively. Particularly, we first calculate the power P_r^i of each subcarrier i for individual receiving antenna r using:

$$P_r^i = \frac{1}{T_e - T_s}\int_{T_s}^{T_e} \left(\left|H_i^{t\leftrightarrow r}(u)\right|\right)^2 \, \mathrm{d}u, \tag{5}$$

where T_s is the starting point of the time window and T_e is the ending point of time window. Given two receiving antennas, r_1 and r_2, we can calculate their power ratio with:

$$PowerRatio_i = \frac{P_{r_1}^i}{P_{r_2}^i}. \tag{6}$$

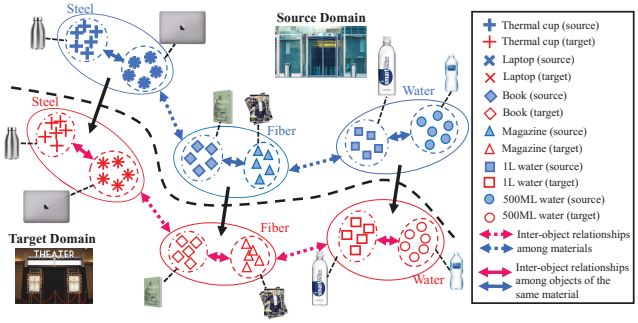


Fig. 5. Illustration of our domain adaptation framework which aligns the material characteristics of the target domain and the source domain.

Since CSI provides measurements of 56 subcarriers, we have 56 power ratio features. Figure 4 (a) shows the power ratio (subcarrier 1, 20, and 40) of different objects. We observe that metal objects (i.e., thermal cup and laptop), fiber objects (i.e., textbook and magazine), and water bottles (i.e., 500ml and 1 L) form different clusters, though the metal objects cannot be well separated from each other.

C. CSI Complex Difference Feature Extraction

We derive CSI complex difference from two receiving antennas based on Equation 1. Based on the magnitude and angle of the CSI complex difference, we calculate 7 statistic features, including mean, variance, median, skewness, kurtosis, IQR, and range. These statistic features capture different perspectives of the signal patterns associated with the object's material and shape. Particularly, we calculate the features for all 56 subcarriers with different operating frequencies and responses to the target object. We extract the 7 statistics features from short frames of CSI, and in total, we extract $1,568$ complex difference features (i.e., from 56 subcarriers). We show the CSI complex difference features of two metal objects (i.e., thermal cup and laptop) in Figure 4 (b). We can observe that the features form two distinctive clusters for the two metal objects, demonstrating their potential to differentiate objects of the same material.

VI. Environment-Independent Object Identification

A. Model Overview

In this work, we develop deep-learning-based models to derive effective representations of objects and enable reliable classification in real-world environments. We designed an *Object Representation Extractor* using convolutional neural network (CNN) to map the CSI complex difference and polarization features into a set of object representations. Then, an *Object Recognizer* is developed to identify objects based on the derived object representation. To mitigate domain impacts and enable environment-independent in-baggage object identification, we develop a *Domain Adaptation Framework* using adversarial training [24]. We refer to different environments as domains. Based on an in-baggage object identification model trained in one domain (i.e., source domain), our domain adaptation framework aims at adapting the existing model to

a new domain (i.e., target domain) using significantly less training data and efforts. We note this is a practical problem setting because we can generate a comprehensive dataset in the lab environment and deploy the system to various public places where engineers only need to retrain the model with a much smaller dataset (e.g., one object of each type of material).

Toward this end, we study the object representations corresponding to different objects. We find that the distances between different objects have similar patterns in the representation space across different domains. Such distances, defined as inter-object relationships, reflect the connections between the objects in terms of their materials and shapes, which do not change in different domains. Specifically, we categorize inter-object relationships into two categories: (1) the relationships among objects of different materials, and (2) the relationships among objects of the same material. We find that the distances between objects of the same material are closer compared to those of different materials, which can be attributed to the greater impacts of the object's material on WiFi signals compared to the object's shape. Figure 5 illustrates the inter-object relationship and the basic idea of our domain adaptation framework.

Based on the inter-object relationships, we develop a domain adaptation framework consisting of two components: the *Material-specific Domain Discriminator* that leverages the relationships among objects of different materials to transfer the material characteristics as a whole from the source domain to the target domain, and the *Inter-object Relationship Alignment* that transfers the shape characteristics learned from the source domain to the target domain by aligning the distributions of object representations of the same material between two domains. These two components allow our system to retrain the in-baggage object identification model using much fewer data from the objects in the target domain. The flow of our domain adaptation framework is shown in Figure 6.

B. CNN-based Object Representation Extractor

We develop a convolutional neural network (CNN) to extract object representations. The input dimension of the CNN model is $(56, 8)$, which involves 7 CSI complex difference statistics and 1 polarization feature from each of the 56 subcarriers. We first use a batch normalization layer to perform input normalization, which reduces small-scale sample inconsistency by removing mean and scaling the samples to unit variance. Three convolutional layers are then used to perform 2D convolution and calculate the 2D representations, which constructively integrates all input features. We use kernel size of $(3, 3)$ for all three convolutional layers, and the strides are all $(2, 1)$. The number of filters for the 3 convolutional layers are 64, 128, 256, respectively. A dropout layer is attached to each convolutional layer to prevent over-fitting. The 2D representations from the third convolution layer is flattened and compressed by a fully connected layer with 128 neurons. By taking the input features X, the CNN model maps the input into object representations Z as follows:

$$Z = CNN(X, \Theta), \tag{7}$$

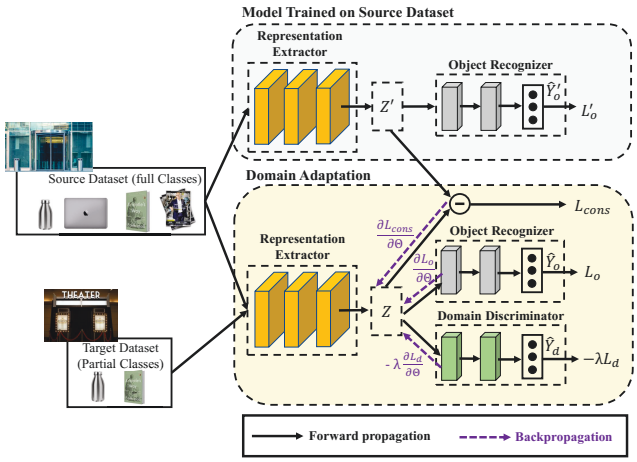


Fig. 6. Flow of our domain adaptation framework.

where Θ represents a set of learnable parameters in the representation extractor. The activations of convolutional layers and the fully-connected layer are all Leaky ReLU.

Based on the extracted object representations Z, we use a neural network with two fully-connected layers to perform object identification. The number of neurons for the two layers are 256 and the activations are Leaky ReLU. We then use a SoftMax layer to predict the object type. Given the input representations Z, the mapping function to the object label $\hat{Y}_o$ is defined as:

$$\hat{Y}_o = G(Z; \Phi), \tag{8}$$

where $G(\cdot)$ represents the mapping function and Φ is the learnable parameters in the neural network. To optimize the object recognizer, we minimize the difference between the groundtruth labels Y_o and the predicted labels $\hat{Y}_o$:

$$L_o = L_{cls}(Y_o, \hat{Y}_o), \tag{9}$$

where $L_{cls}(\cdot, \cdot)$ represents the cross-entropy loss function.

C. Material-Specific Domain Discriminator

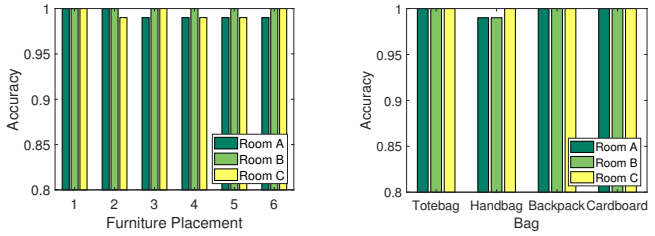
We design a domain discriminator to transfer the material characteristics as a whole from the source domain to the target domain. Particularly, we exploit domain adversarial training [24] to force the object representations of the same material to be similar in the source and target domains. The developed material-specific domain discriminator allows our system to use a small amount of data from the target domain (e.g, one object of each type of material) to achieve the domain adaptation.

As shown in Figure 6, by taking the object representations (i.e., representing one object in the target domain and a few objects of the same material in the source domain) as input (denoted as Z), the domain discriminator predicts the domain labels (i.e., belong to the source domain or target domain) of input samples (denoted as $\hat{Y}_d$). We use the same structure and loss function of the object recognizer (i.e., introduced in Section VI-B) for the domain discriminator, with the domain loss defined as L_d. Predicting the domain labels seems to contradict with our goal of learning domain-independent representations. However, by applying a generative adversarial loss [24], we

(a) Experimental setup at three different room environments

(b) Three bags and one box for the experiments

Fig. 7. Bags and environments for the experiment.



(a) Different furniture placement

(b) Different bags

Fig. 8. The performance of object classification in same domain for different furniture placement settings and different bags.

can use the domain discriminator to guide the representation extractor on learning domain-independent representations, making the source and target representations of the same material follow a similar distribution. The idea is to apply a negative factor $-\lambda$ to the domain loss, L_d, when optimizing the representation extractor, so that the representation extractor is trained to "confuse" the domain discriminator, rendering the extracted representations indistinguishable between the source and the target domain. The loss function for the representation extractor is defined as:

$$L_f = L_o - \lambda L_d, \tag{10}$$

where L_o is the object classification loss. The factor λ is used to control the trade-off between the transferability and the distinctiveness of the learned features.

D. Inter-Object Relationship Alignment

Although the domain discriminator can transfer the objects' material characteristics from the source domain to the target domain, we still need to transfer objects' shape characteristics to the target domain to differentiate objects of the same material. Particularly, we propose to leverage the inter-object relationships among objects of the same material in the source domain, by reusing the object representation extractor trained on the source dataset to align the object representations in the target domain.

Particularly, our alignment method first leverages the representation extractor trained on the source dataset to extract pre-trained representations of all source samples, which encode distinctive shape characteristics for the objects of the same material. Then, we train a new representation extractor and an object recognizer by involving both the source and the target datasets (i.e., all objects in the source dataset and one object per material in the target dataset). By using the pre-trained representations as references, our objective is to make the source representations extracted with the new representation extractor as close as possible to the pre-trained representations, so that the shape characteristics can be transferred to the target domain. We quantify the consistency between source representations (i.e., Z) and the same set of pre-trained representations (i.e., Z') using the mean square error loss:

$$L_{cons} = \frac{1}{K}\sum_{k=1}^{K}(Z_k - Z'_k)^2, \tag{11}$$

where K is the representation dimension. By involving the consistency loss, we modify our adversarial loss and derive an overall loss function to optimize the representation extractor:

$$L_f = L_o + \alpha L_{cons} - \lambda L_d, \tag{12}$$

where the parameter α is used to balance the representation consistency and the transferability. The modified adversarial loss will force the extracted source object representations similar to the pre-trained object representations, so as to transfer the shape characteristics, while also aligning material characteristics of the source and the target domains.

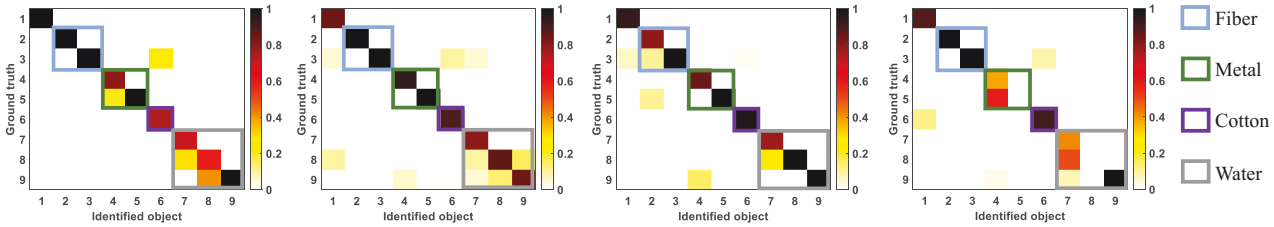
VII. Performance Evaluation

A. Experiment Setup and Methodology

Devices. We use a pair of transmitter and receiver to capture WiFi signals reflected from the object. The transmitter is a directional antenna (i.e., Aaronia HyperLOG 7060) which transmits p-polarized WiFi signals, and the receiver is a pair of orthogonal antennas (i.e., Hawking HD9DP) to capture the signal polarization. The transmitter and receiver are connected to two desktops installed with Atheros WiFi network interface cards (i.e., Atheros QCA9590), which report CSI of 56 subcarriers. We extract the polarization and CSI complex difference features based on the CSI from the two orthogonal receiving antennas. The WiFi packet transmission rate is set to 1000 $pkts/s$.

Data Collection. Experiments are conducted in three different university offices as shown in Figure 7 (a), and the room sizes are $15ft \times 13ft$, $16ft \times 12ft$, and $28ft \times 23ft$, respectively. We collect data of 14 representative objects of four different materials including *fiber:* book, magazine, newspaper; *metal:* thermal cup, laptop; *cotton/polyester:* cotton T-shirts ($\times 2$), cotton T-shirts ($\times 4$), hoodie, polyester T-shirts, polyester pants; *water:* 1L bottle with 1L water, 1L bottle with 500ML water, 500ML bottle with 500ML water. To evaluate our system on in-baggage object identification, we collect data with objects placed in three different bags and one box as shown in Figure 7 (b). Furthermore, for each room environment, we collected data on different days and with different furniture settings (i.e., two desks and two chairs were moved at least 3ft).

Evaluation Methodology. We separate the collected data into source datasets and target datasets with the bags/furniture settings/days/rooms. We use half of the data of the selected objects in the target dataset for domain adaptation. We define the object identification accuracy as the percentage of predicted objects that are correctly recognized.



(a) Cross different bags (b) Cross different furniture settings (c) Cross different days (d) Cross different rooms

Fig. 9. Confusion matrix of object identification accuracy under various real-world dynamics.

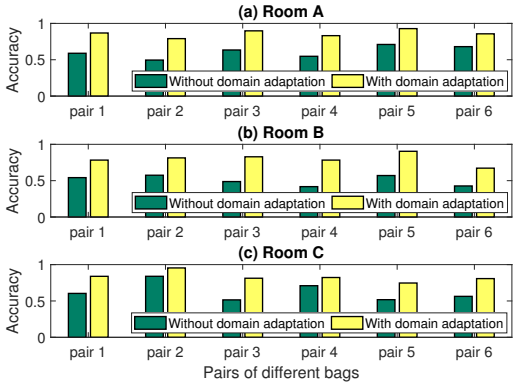


Fig. 10. Performance of object identification across different bags. We choose six bag pairs involving all four bags for the evaluation.

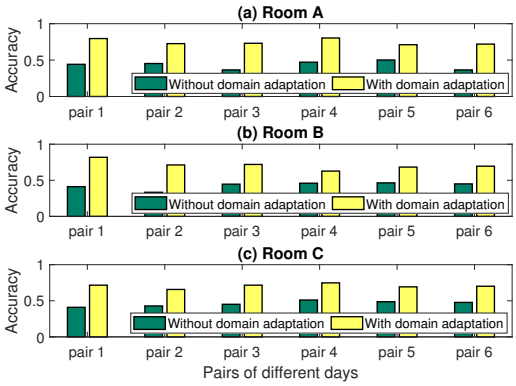


Fig. 11. Performance of object identification across different days.

B. Performance of Object Identification

We first evaluate the performance of our object identification system in the same environment, by randomly splitting training and testing sets. Figure 8 (a) shows that the object identification accuracy for individual furniture settings. We find that our system achieves over 97% accuracies for all three room environments. The results demonstrate the effectiveness of our system for object identification. We also evaluate the performance of our system with different bags. Specifically, four types of representative bags, including a tote bag, a handbag, a backpack, and a cardboard box, are used to enclose the target objects. Figure 8 (b) shows that our system can achieve above 99% accuracy for all four bags in different rooms, which shows the effectiveness of the proposed system.

C. Performance of Bag-independent Object Identification

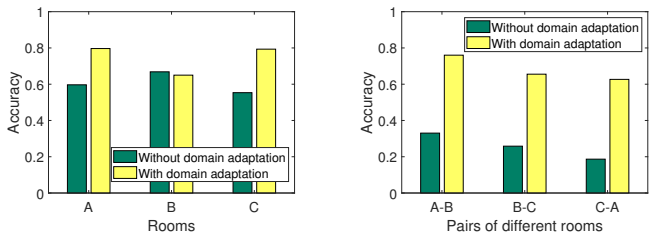
We then evaluate the performance of domain adaptation cross different bags, by using data collected with one bag as the source dataset and those collected with a different bag as a target dataset. In total, six bag pairs are considered for domain adaptation for each room environment. We train a deep learning model (i.e., a feature extractor and an object recognizer) without applying domain adaptation as the baseline model. The object identification accuracy for individual bag pair are shown in Figure 10. We find that without domain adaptation, the object identification accuracies are less than 70% for all three rooms. By applying the proposed domain adaptation framework, we find that our system can achieve over 83% accuracies on average. We also find that the domain adaptation performance is slightly different for different bag pairs, which can be attributed to the difficulties for adaptation. For all bag pairs, our system can achieve more than 78% accuracies, which demonstrates the effectiveness of the proposed domain adaptation framework.

D. Performance of Time-independent Object Identification

To evaluate our system across days, we use data collected on one day as the source domain and those collected on a different day as the target domain. We keep the same furniture setting when collecting the data. We collect data from three days in each room environment and there are six pairs of days in total. Figure 11 shows the object identification performance under different day pairs for all three rooms. We find that without the proposed domain adaptation, the object identification accuracies are only around 50%, but with the domain adaptation, our system can achieve over 72% accuracies on average for all pairs of days (i.e., 22% improvement). Interestingly, we find that the accuracy is lower compared to domain adaptation across bags. Such results show that combating ambient interference is more challenging than mitigating the impacts of bags. In general, the results show the proposed domain adaptation framework is effective in mitigating time-varying ambient interference.

E. Performance of Environment-independent Object Identification

Small-scale Furniture Variations. To evaluate performance domain adaptation across different furniture settings, we use data of one furniture setting as the source domain and data of a different furniture setting as the target domain. For each room environment, we collect data of three different furniture placement settings and we have six pairs of settings in total. Figure 12 (a) shows the average results for each room. We find that without domain adaptation, the object identification accuracy with most furniture placement pairs in the three



(a) Different furniture settings (b) Different rooms

Fig. 12. Performance of object identification across different environments, including different furniture placement settings and rooms.

rooms are below 60%. After applying our adaptation method, the performance improved by over 20%, and the average accuracies at room A and C reach 80%. The results show that the proposed domain adaptation method well improve object identification performance across room environments.

Large-scale Room Changes. Finally, we use data collected in one room as the source domain and the data collected in another room as the target domain. Such an evaluation setting is the most challenging since the WiFi signal patterns will be very significant different due to different room layouts. For each room, we collect data on three different days. Figure 12 (b) shows the object identification performance when using room A as the source domain and room B as the target domain. Our system can achieve 79.1% accuracy, which 36.2% improvement compared to the accuracy without domain adaptation. Figure 9 (d) shows the confusion matrix of object identification for the room pair $A-B$. We find that even for some objects that are not correctly classified, they can still be detected as belonging to the correct material. We also find that even for the room pairs (i.e., $B-C$, $C-A$), where the model only has less than 30% accuracy, our domain adaptation framework can still improve the performance to over 65%. The significant improvements show the effectiveness of the proposed domain adaptation framework.

VIII. Conclusion

In this paper, we propose an environment-independent object identification system by leveraging low-cost WiFi. We find that the polarization features and CSI complex difference statistics can well capture the material and shape characteristics of in-baggage objects. To mitigate the impacts of real-world dynamics including different types of bags, time-varying ambient interference, and physical environment changes, we develop a deep learning model based on CNN to learn effective and robust abstract representations of objects. Furthermore, we design a domain adaptation framework to learn and leverage the inter-object relationships from the source domain to reduce the training efforts in the target domain, which significantly reduce the training cost during deployment. The experimental results demonstrate the effectiveness and generalizability of the proposed system.

IX. Acknowledgment

This work was partially supported by the National Science Foundation Grants CCF-1909963, CCF-2000480, CCF-2028876, CCF-2028894, CNS-1566455, CNS-1717356, CNS-1820624, CNS-1815908, ECCS-2033433.

References

[1] T. N. Y. Times, "A list of mass shootings in the united states in 2021," https://www.nytimes.com/article/mass-shootings-2021.html, 2021.

[2] statista, "Number of fatalities in explosion incidents in the united states from 2012 to 2019," https://www.statista.com/statistics/785966/number-of-fatalities-in-explosion-incidents-in-the-united-states/, 2021.

[3] D. Turcsany, A. Mouton, and T. P. Breckon, "Improving feature-based object recognition for x-ray baggage security screening using primed visualwords," in *proc. of IEEE ICIT*, 2013, pp. 1140–1145.

[4] G. T. Flitton, T. P. Breckon, and N. M. Bouallagu, "Object recognition using 3d sift in complex ct volumes." in *BMVC*, vol. 1, 2010, pp. 1–12.

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

[6] B. Xie, J. Xiong, X. Chen, E. Chai, L. Li, Z. Tang, and D. Fang, "Tagtag: material sensing with commodity rfid," in *proc. of ACM SenSys*, 2019, pp. 338–350.

[7] H.-S. Yeo, G. Flamich, P. Schrempf, D. Harris-Birtill, and A. Quigley, "Radarcat: Radar categorization for input & interaction," in *proc. of ACM UIST*, 2016, pp. 833–841.

[8] Y. Zhu, Y. Zhu, B. Y. Zhao, and H. Zheng, "Reusing 60ghz radios for mobile radar imaging," in *proc. of ACM MobiCom*, 2015, pp. 103–116.

[9] C. Li, Z. Liu, Y. Yao, Z. Cao, M. Zhang, and Y. Liu, "Wi-fi see it all: generative adversarial network-augmented versatile wi-fi imaging," in *proc. of ACM SenSys*, 2020, pp. 436–448.

[10] J. Qiu, D. Chu, X. Meng, and T. Moscibroda, "On the feasibility of real-time phone-to-phone 3d localization," in *proc. of ACM SenSys*, 2011.

[11] C. Wang, J. Liu, Y. Chen, H. Liu, and Y. Wang, "Towards in-baggage suspicious object detection using commodity wifi," in *proc. of IEEE CNS*, 2018, pp. 1–9.

[12] J. Ševek, S. Katuščák, L. Dubinyová, and M. Fikar, "An automatic identification of wood materials from color images," in *proc. of IEEE K&I*, 2016, pp. 1–6.

[13] K. Tanaka, Y. Mukaigawa, T. Funatomi, H. Kubo, Y. Matsushita, and Y. Yagi, "Material classification from time-of-flight distortions," *IEEE transactions on pattern analysis and machine intelligence*, vol. 41, no. 12, pp. 2906–2918, 2018.

[14] W. Mao, M. Wang, and L. Qiu, "Aim: Acoustic imaging on a mobile," in *proc. of ACM MobiSys*, 2018, pp. 468–481.

[15] A. Dhekne, M. Gowda, Y. Zhao, H. Hassanieh, and R. R. Choudhury, "Liquid: A wireless liquid identifier," in *proc. of ACM MobiSys*, 2018, pp. 442–454.

[16] Y. Wang, J. Liu, Y. Chen, M. Gruteser, J. Yang, and H. Liu, "E-eyes: device-free location-oriented activity identification using fine-grained wifi signatures," in *proc. of ACM MobiCom*, 2014, pp. 617–628.

[17] J. Liu, Y. Wang, Y. Chen, J. Yang, X. Chen, and J. Cheng, "Tracking vital signs during sleep leveraging off-the-shelf wifi," in *proc. of ACM MobiHoc*, 2015, pp. 267–276.

[18] D. Huang, R. Nandakumar, and S. Gollakota, "Feasibility and limits of wi-fi imaging," in *proc. of ACM SenSys*, 2014, pp. 266–279.

[19] C. Feng, J. Xiong, L. Chang, J. Wang, X. Chen, D. Fang, and Z. Tang, "Wimi: Target material identification with commodity wi-fi devices," in *proc. of IEEE ICDCS*, 2019, pp. 700–710.

[20] C. R. Karanam and Y. Mostofi, "3d through-wall imaging with unmanned aerial vehicles using wifi," in *2017 16th ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN)*. IEEE, 2017, pp. 131–142.

[21] D. Zhang, J. Wang, J. Jang, J. Zhang, and S. Kumar, "On the feasibility of wi-fi based material sensing," in *The 25th Annual International Conference on Mobile Computing and Networking*, 2019, pp. 1–16.

[22] S. Tominaga and A. Kimachi, "Polarization imaging for material classification," *Optical Engineering*, vol. 47, no. 12, p. 123201, 2008.

[23] S. Sen, B. Radunovic, R. R. Choudhury, and T. Minka, "You are facing the mona lisa: Spot localization using phy layer information," in *proc. of ACM MobiSys*, 2012, pp. 183–196.

[24] Y. Ganin, E. Ustinova, H. Ajakan, P. Germain, H. Larochelle, F. Laviolette, M. Marchand, and V. Lempitsky, "Domain-adversarial training of neural networks," *The journal of machine learning research*, vol. 17, no. 1, pp. 2096–2030, 2016.