IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 18 August 2025, accepted 1 September 2025, date of publication 8 September 2025,
date of current version 12 September 2025.

Digital Object Identifier 10.1109/ACCESS.2025.3606941

RESEARCH ARTICLE

PoBiTag: Toward Unobtrusive and Customizable Tag-Based Interaction With Optical Polarization and Birefringence

SEOHYEON PARK, SEUNGHUN CHAE, JAEMIN CHOI, AND HYOSU KIM, (Member, IEEE)
School of Computer Science and Engineering, Chung-Ang University, Seoul 06974, South Korea

Corresponding author: Hyosu Kim (hskimhello@cau.ac.kr)

This work was supported in part by the Chung-Ang University Graduate Research Scholarship in 2023; in part by the National Research Foundation of Korea (NRF) Grant Funded by the Korea Government (MSIT) under Grant RS-2025-02217071; and in part by Institute of Information and Communications Technology Planning and Evaluation (IITP) Grant Funded by the Korea Government (MSIT) under Grant RS-2025-02305436. Development of Digital Innovative Element Technologies for Rapid Prediction of Potential Complex Disasters and Continuous Disaster Prevention.

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Institutional Review Board of Chung-Ang University.

ABSTRACT Traditional tag-based interaction systems often rely on visible codes such as QR codes or barcodes, which can disrupt the visual design of printed media, particularly in content-dense or space-limited environments. This work introduces PoBiTag, a visually unobtrusive tagging system that uses common transparent and birefringent materials (e.g., PET stickers or cellophane tape) as tags. This tag design allows users to embed tags on any surface of interest without compromising their visual aesthetics. To enable accurate and convenient tag scanning, PoBiTag enhances tag visibility based on optical phenomena including polarization and birefringence. It converts a mobile device's camera and torch into polarized optical components by attaching polarizing films. This enables illumination and capture of tagged surfaces under polarized light, making birefringent regions visually distinguishable from others and thus allowing accurate tag segmentation and identification. Experiments with real-world users demonstrate that PoBiTag significantly reduces tag-scanning time, even on complex or glossy surfaces, while achieving a high tag identification accuracy of 96.5% on average. These results confirm PoBiTag's practicality for diverse real-world tag-based interaction scenarios, such as interactive books.

INDEX TERMS Unobtrusive tags, tag-based interaction, polarization, birefringence.

I. INTRODUCTION

Optical tags, such as QR codes [1], [2] and barcodes [3], have been widely adopted in the real world to bridge the gap between the physical and digital worlds. They enable smartphone-based payments in retail stores, provide instant access to web content from printed materials, and facilitate contactless check-ins at events or transport hubs.

Yet, optical tags face several limitations that hinder their applicability, especially in space-constrained or content-dense scenarios. For example, interactive books enhance the reading experience by allowing users to access audio, video, or augmented reality content through tag-based interactions. However, incorporating multiple tags within limited page space can disrupt visual design and reduce room for primary content, making it difficult to scale such interactions without compromising aesthetics. Furthermore, customizing tag-based interaction after printing is also challenging. Users might want to add new tags to link to additional multimedia content, but doing so often results in overlap with existing content, thereby limiting flexibility after production. As an alternative, infrared (IR)-based tagging system [4], [5], [6], [7], [8], [9], [10] has been proposed. They preserve visual aesthetics using invisible IR tags, but require specialized hardware (e.g., IR cameras).

This paper presents a novel unobtrusive and virtually invisible tagging system, called *PoBiTag*, that leverages common transparent stickers (or tapes) made of PET or

The associate editor coordinating the review of this manuscript and approving it for publication was Hengyong Yu.

© 2025 The Authors. This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License.
For more information, see https://creativecommons.org/licenses/by-nc-nd/4.0/



FIGURE 1. Usage scenario of PoBiTag.

cellophane for tag-based interactions. As illustrated in Figure 1, users with PoBiTag define interactions by *i*) attaching a transparent sticker to a surface of interest (e.g., printed materials), *ii*) capturing an image of the tagged area, and *iii*) mapping a desired functionality (e.g., multimedia content playback) to the tag. In most scenarios, users are already aware of the presence of the tag and have a rough idea of where it is attached. Because the tag is almost invisible, it is difficult to clearly locate or recognize it with the naked eye. To address this, once registered, PoBiTag allows users to scan the general area to reveal the tag through its enhanced camera view, enabling precise selection and launching of the associated functionality. Specifically, PoBiTag determines which interaction to trigger by comparing the selected tag area with a set of pre-registered tag images.

Such a tagging system based on transparent stickers comes with several benefits. First, because the tags are visually unobtrusive, they can be seamlessly integrated into existing designs without disrupting the visual aesthetics of the surface. That is, it allows users to attach tags freely to any region of interest. Second, since the transparent sticker overlays the underlying content, each tag naturally exhibits a unique appearance based on its background without requiring additional tag design effort. However, a major challenge remains in enabling tag-based interactions with transparent stickers. Unlike conventional tags with distinct visual boundaries, transparent stickers blend into the background, making it difficult to accurately isolate and detect them. Even worse, PoBiTag must ensure robust and consistent detection across diverse surfaces, printing materials, and lighting conditions.

To address this challenge, a tag detection method is designed that can be applied to any commercial mobile device (e.g., smartphones and tablets), based on optical phenomena, called *polarization* and *birefringence*, as follows:

- *Optics-based tag visibility enhancement.* To enhance tag visibility, linear polarizing films are attached to both the rear-facing camera and torch, aligned in parallel. When the polarized torch emits light, it reflects off background surfaces and returns to the camera. For non-birefringent surfaces, the reflected light preserves its original polarization and passes through the camera's polarizer, resulting in a bright appearance. In contrast, when the light reflects from a birefringent material, such as a PET or cellophane sticker, its polarization state is altered due to the phase shift introduced by birefringence. This creates strong visual contrast, making birefringent regions appear darker than the background.
- *Tag identification with minimal user involvement.* The enhanced view, which visually distinguishes tags from the background, is shown on the screen, allowing users to easily detect and tap potential tags from it. Upon selection, PoBiTag captures the corresponding frame and segments the tag using the slim Segment Anything Model (SlimSAM) [11]. It then identifies the tag via a deep learning-based classifier trained on labeled samples collected during registration.

This study implements a prototype of PoBiTag, which consists of a camera accessory for tag visibility enhancement and software algorithms for tag segmentation and identification. Extensive experiments with the prototype implementation demonstrate that PoBiTag significantly improves the visibility of tag and reduces the time required to detect tags on various types of surfaces. It also achieves high tag recognition accuracy (e.g., 96.5% on average) with negligible training effort. PoBiTag maintains robust performance across diverse environments and situations, including device orientations and lighting conditions.

The contributions can be summarized as follows:

- PoBiTag is proposed as a novel tagging system that leverages optical phenomena, polarization and birefringence, to enable unobtrusive and customizable tag-based interaction.
- Hardware and software tagging solutions that can be easily deployed in diverse real-world environments. Specifically, cellophane tape or PET stickers are used as tags, making them easily detectable under polarized imaging while remaining virtually invisible to the naked eye.
- PoBiTag is evaluated through extensive experiments under varying lighting conditions and angles, demonstrating its effectiveness, robustness, and practicality for real-world use.

II. RELATED WORKS

This section introduces previous optical tagging systems, which associate digital information with physical objects using tags, and clarify how our work differs from them.

A. VISIBLE TAGGING SYSTEM

Visible tags, which can be recognized by the human eye, have been widely used due to their easy fabrication and straightforward visual verification without the need for specialized equipment. The most representative examples are 2D barcodes [3] and QR codes [1], [2] used in diverse real-world

scenarios, such as payment, authentication, and access to information. Furthermore, many researchers have proposed alternative visible tagging approaches to support diverse application needs. For instance, fiducial markers, such as AR Tag [12], AprilTag [13], and CalTag [14] use simple black-and-white patterns to enable precise localization, making them popular in robotics and augmented reality applications. AniCode [15] allows printed codes to trigger animations when captured by a smartphone camera, introducing a layer of interactivity. ClayCode [16] uses a tree-structured arrangement of color regions embedded within polygons, achieving high recognition accuracy from various angles and surfaces while also supporting artistic expression.

Despite their wide adoption, visible tags inherently interfere with aesthetic design or content layout, potentially causing visual clutter or limiting usable space. In contrast, PoBiTag employs transparent stickers (or tapes) as tags, allowing them to remain virtually invisible to the naked eye and thus preserve the visual aesthetics of tagged surfaces.

B. INVISIBLE OR UNOBTRUSIVE TAGGING SYSTEM

Similarly to PoBiTag, there have been several attempts to overcome the limitations of visible tagging systems by embedding digital information into invisible or unobtrusive tags.

1) POLARIZATION-BASED TAGS.

PolarTag [17] constructs transparent tags by combining a polarizing film with cellophane tapes. Specifically, it stacks multiple layers of cellophane tape onto the polarizing film to generate unique optical patterns that can be detected using a polarized camera. In this design, the encoded patterns remain invisible under normal lighting conditions but become recognizable under controlled polarization setups. However, since the polarizing film itself appears visually dark or tinted, the physical presence and shape of the tag remain noticeable to the human eye, which can negatively affect the visual aesthetics and unobtrusiveness of the tagged surface.

2) INFRARED (IR)-BASED TAGS.

IR-based tagging systems leverage invisible infrared light to embed information. For example, LayerCode [5] and InfraredTags [4] integrate infrared materials, such as near-infrared (NIR) dyes or IR-transmissive filament, into 3D-printed objects, to embed markers that can be recognized by NIR or low-cost IR cameras. Imprinto [7] prints infrared tags on paper using standard inkjet printers with IR-absorptive inks. HideOut [8] combines IR-absorbing markers and a projector-camera setup for interactive content projection. MiniStudio [9] tracks miniature objects by attaching IR-reflective stickers, while Spyn [10] maps digital data to fabric surfaces using IR ink for precise tracking with IR cameras. The invisibility of IR light allows such IR-based approaches to provide an unobtrusive tag-based interaction. However, they often require specialized hardware, such as IR cameras, which limits their applicability on commodity mobile devices.

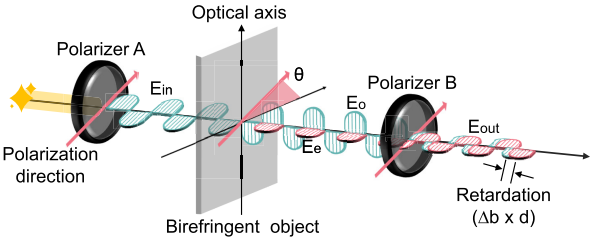


FIGURE 2. Visible intensity (or color pattern) variation of polarized light induced by birefringent materials.

3) STRUCTURALLY EMBEDDED TAGS.

Some works [6], [18], [19], [20] encode information invisibly by adding microstructures or air gaps inside objects. Specifically, AirCode [18] inserts internal air gaps of unique patterns during 3D printing and recognizes the patterns using both a light projector and camera. InfraStructs [6] employs terahertz scanning to retrieve data from air-layer-based internal structures. G-ID [19] modifies slicing parameters during 3D printing to create nearly invisible surface textures that encode identifiers without affecting object design. Seedmarkers [20] fabricates Voronoi diagram-based marker patterns via 3D printing or laser cutting and integrates them into manufactured objects. Another class of works [21], [22], [23], [24] identify objects or recognize tags based on their unique acoustic structures that produce sound with specific frequencies. For example, Acoustic Barcode [24] extracts information from unique sound patterns generated when surface notches, fabricated by laser cutting, are scraped. Others [21], [22], [23] add resonant cavities or filters within objects, enabling electronic-free identification using microphones. However, these structurally embedded approaches can embed tags only during the object manufacturing process (e.g., 3D printing or fabrication), limiting their applicability for post-production tagging or user-level customization.

Unlike these previous works, PoBiTag provides a user-customizable, easily deployable, and unobtrusive tagging solution by *i*) utilizing transparent sticker or tapes as tags to conceal information without visual disturbance and *ii*) proposing tag identification methods that require only a camera and torch, both of which are readily available on commodity mobile devices.

III. BACKGROUND: POLARIZATION AND BIREFRINGENCE

In this section, two key optical phenomena, polarization and birefringence, on which PoBiTag is based, are described.

Polarization, in the context of electromagnetic waves such as light, refers to the orientation of the oscillations of the wave's electric field. Most natural and artificial light sources, such as sunlight or bulbs, emit unpolarized light in which the electric field oscillates in all directions perpendicular to the direction of propagation. However, when light is reflected, refracted, or scattered, the oscillations of its electric field can become biased toward specific directions, resulting in partially polarized light. Polarization can also be controlled

using optical filters, called polarizers. A common example is a linear polarizer, which only allows light waves vibrating in a specific direction to pass through. This filtering effect occurs because the internal structure of the polarizer absorbs or blocks components of the electric field that oscillates in directions orthogonal to its polarization axis. As a result, the light passing through the polarizer is uniformly polarized with its electric field oscillating along the direction defined by the polarizer. On the basis of this principle, polarizers are widely used in various real-world applications, e.g., polarized sunglasses and liquid crystal display (LCD) screens.

Birefringence is an optical property of anisotropic materials such as transparent plastics, crystals, or polymers. When light passes through these materials, the materials split the light into two orthogonal polarization components, commonly referred to as the ordinary and extraordinary rays. These components experience different refractive indices and thus propagate at different speeds, introducing a phase difference $\Delta\phi$ between them:

$$\Delta\phi = \frac{2\pi}{\lambda} \cdot \Delta n \cdot d, \tag{1}$$

where λ is the wavelength of light in vacuum, Δn is the birefringence (difference in refractive indices), and d is the thickness of the birefringent material.

If the incident light is polarized, the phase difference introduced by the birefringent material alters the polarization state of the outgoing light. In addition, when this modified light passes through another polarizer, it produces visible intensity variations (or color patterns). As illustrated in Figure 2, suppose that a birefringent material is placed between two parallel polarizers with an angle θ between the material's optical axis and the polarization direction of the polarizers. When the light polarized by the first polarizer encounters the material, its electric field (E_{in}) is decomposed into two orthogonal components:

$$\begin{aligned} E_o &= E_{in} \sin\theta, \\ E_e &= E_{in} \cos\theta, \end{aligned} \tag{2}$$

where E_o and E_e are aligned perpendicularly to and along the optical axis, respectively. After propagating through the material, these two components have a phase difference $\Delta\phi$, forming the resultant electric field E_b as:

$$E_b = E_o + E_e \cdot e^{i\Delta\phi}. \tag{3}$$

The altered electric field is then projected onto the polarization direction of the second polarizer, resulting in the outgoing electric field E_{out} as

$$\begin{aligned} E_{out} &= E_o \sin\theta + E_e \cdot e^{i\Delta\phi} \cdot \cos\theta \\ &= sE_{in}(\sin^2\theta + \cos^2\theta \cdot e^{i\Delta\phi}). \end{aligned} \tag{4}$$

Thus, the intensity of the outgoing light $|E_{out}|^2$ differs from that of the incident light $|E_{in}|^2$ unless θ is equal to 0 or $\pi/2$, where the polarization direction is perfectly aligned with or perpendicular to the optical axis of the material, resulting in no interference between the orthogonal components.

IV. PoBiTag DESIGN

PoBiTag is an unobtrusive and customizable tagging system built on the optical phenomena of polarization and birefringence. Specifically, it utilizes transparent birefringent stickers (or tapes) as tags and detects them by collaboratively using a mobile device's built-in camera and torch, each covered with linear polarizing films. This optical setup improves tag visibility to help users easily identify the tags. Figure 3 illustrates the overall workflow of PoBiTag to enable tag-based interaction under this optical setup. It consists of two phases, *registration* and *interaction*:

- *Tag registration.* In the registration phase, users define a new tag interaction through three simple steps. First, they attach a tag, i.e., a transparent sticker, to a surface of interest (e.g., printed materials). Second, they capture an image of the tagged area using their mobile device's camera. This image serves as the visual signature of the tag for training a tag classification model. Third, they associate a desired functionality, such as playing multimedia content or launching an application, with the tag.
- *Tag interaction.* In the interaction phase, users point their mobile device at a surface containing one or more registered tags. At this time, PoBiTag illuminates the scene using the polarized torch and captures it using the polarized camera. The resulting images, where birefringent regions become visually distinct, are then displayed in real time on the device's screen, allowing users to easily select a tag of interest. When a tag is tapped, PoBiTag segments the selected region from the corresponding frame, uses the cropped image to identify the tag via the pre-trained classifier, and finally triggers the functionality mapped to the recognized tag.

Through this process, PoBiTag enables seamless tag-based interaction using tags that are virtually invisible on real-world surfaces, yet easily detectable through commodity mobile devices. In the remainder of this section, we present the design of each key enabling technique in detail.

A. OPTICS-BASED TAG VISIBILITY ENHANCEMENT

Transparent stickers allow users to embed tag-based interactions into any surface of interest while preserving its visual aesthetics. However, this transparency inherently introduces a significant challenge: the tags become difficult to visually detect. Even if users are already aware of the presence of tags from the registration phase, it is still difficult to spot their exact positions. To address this, PoBiTag enhances the visibility of transparent tags in camera-captured recordings by leveraging the optical phenomena where the polarization state of light changes when it passes through birefringent materials.

Specifically, PoBiTag employs transparent stickers made from birefringent materials as tags. These materials are selected based on the following criteria. First, they should

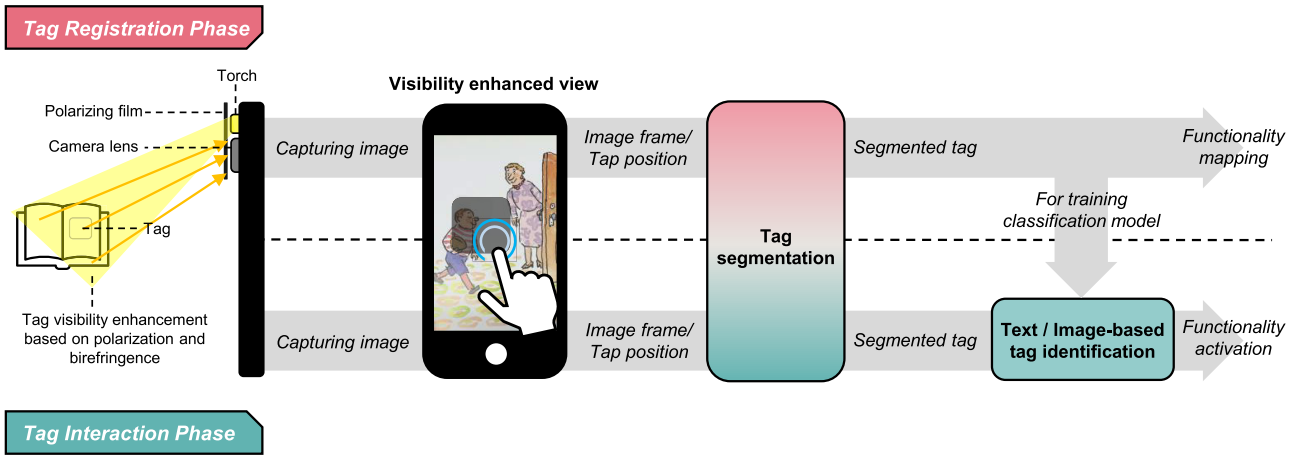


FIGURE 3. Overall architecture of PoBiTag.

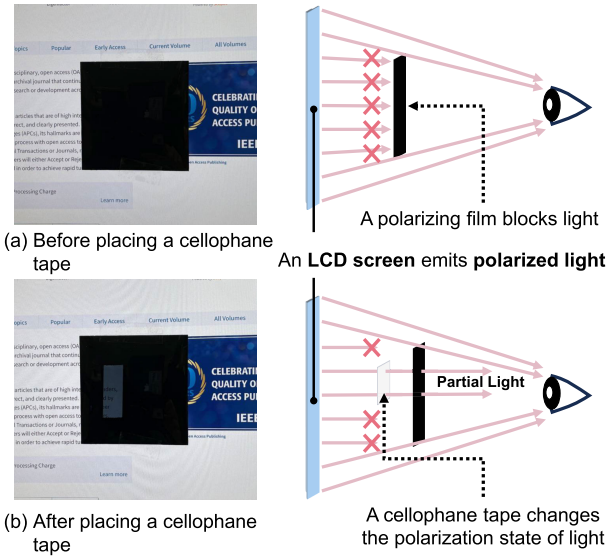


FIGURE 4. Effect of birefringent materials on polarized light. We observed the effect by placing a cellophane tape between an LCD screen and a polarizing film perpendicularly aligned with the screen.

induce noticeable changes in the polarization state of light. As explained in Section III, the birefringence of a material (Δn) determines the phase difference ($\Delta\phi$) between the two orthogonal components of light passing through it. Therefore, we use materials with high birefringence that can produce significant alterations in the polarization state of transmitted light. Second, the tag materials should be readily available in everyday life. As presented in Table 1, while materials like calcite exhibit extremely high birefringence, they are not commonly accessible, making them impractical for use as tags. In contrast, materials such as cellophane and PET, are widely used in everyday products like adhesive tape and transparent labels, and offer moderate birefringence. For example, when a piece of cellophane tape is placed on an LCD screen and viewed through a polarizing film oriented

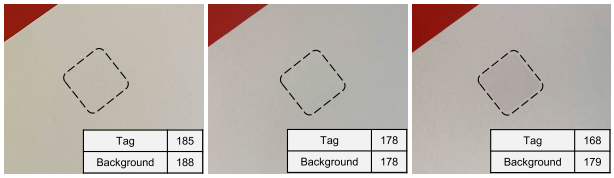
TABLE 1. Common birefringent materials.

Material	Δn	Usage
Calcite ($CaCO_2$)	0.172 [25]	Optics (e.g., prisms)
Cellophane	0.011 [26]	Packaging and adhesive tapes
Polyethylene Terephthalate (PET)	<0.05 [27]	Labels and packaging films
Polycarbonate	0.0006 [28]	Lenses and display panels

perpendicular to the polarization direction of the screen, the region covered by the tape becomes visibly bright (see Figure 4). This is because the cellophane's birefringence modifies the polarization state of the emitted light, allowing it to partially pass through the crossed polarizer. Inspired by this, PoBiTag uses common cellophane shtape or PET stickers as tags, enabling users to embed tags on any surface of interest without the need for specialized or custom-manufactured materials.

The birefringent tags are captured with enhanced visibility using two hardware components: a polarized torch and a polarized camera. PoBiTag first illuminates a target surface (e.g., printed materials containing tags) using the polarized torch. The emitted light reflects off both birefringent and non-birefringent regions, experiencing different alterations in polarization state. For example, light reflected from non-birefringent regions typically retains its original polarization, while light interacting with birefringent tags undergoes a phase shift that modifies its polarization state. PoBiTag then captures these reflections using the polarized camera, whose polarization axis is aligned parallel to that of the torch. It should be noted that the built-in camera and torch of a mobile device can be easily converted into polarized components by attaching external polarizing films. Details of our hardware implementation for this conversion are provided in Section V.

The parallel alignment between the polarized camera and torch is configured based on the following observations.



(a) Non-polarized (b) Polarized and perpendicularly-aligned (c) Polarized and parallelly-aligned

FIGURE 5. Effect of polarization axis alignment on birefringent area visibility. Note that we computed the brightness of birefringent and non-birefringent areas as the mean of the brightness values inside and outside a tag boundary.

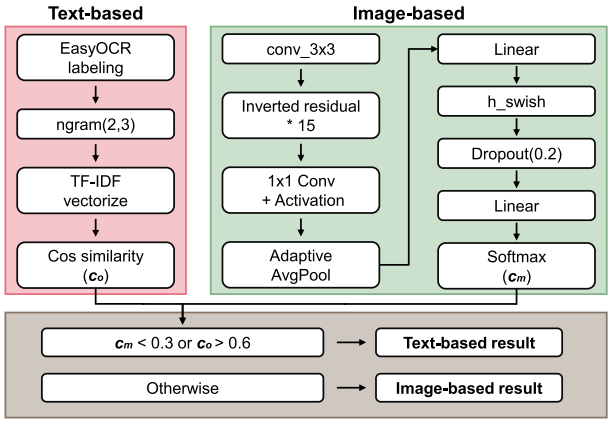
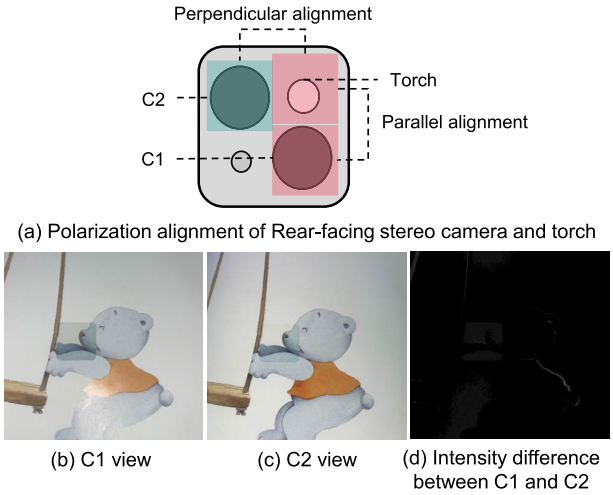


FIGURE 6. Tag identification model structure.

As demonstrated in Figure 5(a), attaching transparent stickers causes the tagged regions to be slightly darker than the surrounding areas, due to subtle differences in light diffusion. When the polarization axes of the camera and torch are set to perpendicular to each other, light reflected from non-birefringent regions are mostly blocked by the polarizing filter of the camera while a portion of light from birefringent tags passes through the filter (see Figure 5(c)). This makes it difficult to distinguish between tagged and non-tagged regions, as both can appear with similar brightness in the captured image. In contrast, aligning the polarization axes in parallel ensures that reflections from non-birefringent surfaces appear bright, while birefringent regions are attenuated due to partial polarization misalignment as explained in Section III. This setup maximizes contrast and enables reliable tag detection (see Figure 5(b)).

PoBiTag further enhances the visibility of tags in the captured images through a two-stage image processing method. First, CLAHE (Contrast Limited Adaptive Histogram Equalization) [29] is applied to the luminance channel in the CIELAB color space. Specifically, the clip limit for CLAHE is set to 0.8 to gently enhance contrast in low-brightness areas (e.g., birefringent regions) and reduce the risk of noise amplification. Second, we adjust the global brightness and contrast of the image through gamma correction with a gamma value of 0.5, which is empirically set to enhance contrast between birefringent and non-birefringent areas.



(a) Polarization alignment of Rear-facing stereo camera and torch

(b) C1 view (c) C2 view (d) Intensity difference between C1 and C2

FIGURE 7. Rear-facing stereo camera and torch with a polarizing film.

B. LIGHTWEIGHT TAG IDENTIFICATION

While a user continuously captures a surface of interest using PoBiTag, visibility-enhanced views are displayed on the device's screen in real-time. Interaction with the tags is triggered via the user's touch inputs. More specifically, when one of the visible tags is tapped, PoBiTag takes the corresponding image frame and the tapped location and exploits them for both tag segmentation and identification as follows.

First, the selected tag area is segmented from the frame computationally efficiently by employing SlimSAM [11], a lightweight segmentation model derived from Segment Anything Model [30] with PyTorch-based pruning techniques [31]. To preserve segmentation accuracy despite the model's reduced complexity, the user's tapped location, captured during interaction, is incorporated as a spatial cue for segmentation. Specifically, a triangle-shaped set of points is generated centered around the tapped position, which represents an approximate area of a fingertip. These points are used as prompts in SlimSAM's inference pipeline, guiding the model to focus on the relevant local region. This spatial prompting enables accurate segmentation of the tag boundary, even on visually complex or low-contrast surfaces. Additionally, to prevent generating excessively large segmentation masks, which are not common in real-world tagging scenarios, we added a mask prompt that constrains the maximum size of the segmented region.

Note that in the cases where dual cameras are available, PoBiTag collaboratively uses them to achieve more accurate tag segmentation. Specifically, as illustrated in Figure 7, polarizing films are attached to each camera lens with different polarization orientations. One camera is polarized parallel to the polarization direction of the torch, while the other camera's polarizer is oriented perpendicular to the torch. Under this configuration, polarized light emitted from the torch reflects off birefringent areas (such as tags)

and is captured by the two cameras with distinct intensity differences, as explained in Section IV-A. Based on this, a pixel-wise brightness difference ($\Delta b(x, y)$) between the two camera frames is computed as follows:

$$\Delta b(x, y) = \begin{cases} 0, & \text{if } b_1(x, y) \geq b_2(x, y) \\ |b_1(x, y) - b_2(x, y)|, & \text{otherwise,} \end{cases} \tag{5}$$

where $b_1(x, y)$ and $b_2(x, y)$ denote the grayscale brightness values at pixel (x, y) for the frames obtained from the parallel- and perpendicular-polarized cameras, respectively. This computation amplifies the contrast between birefringent and non-birefringent regions, as shown in Figure 7. Using $\Delta b(x, y)$ as input, PoBiTag then extracts tag outlines through segmentation and crops the corresponding tag regions from the original image captured by the parallel-polarized camera based on the segmented outlines.

Once the selected area is segmented, PoBiTag identifies a corresponding tag using the cropped image as a visual signature for classification. Figure 6 illustrates the structure of the deep learning model used for tag identification, designed to balance between classification accuracy and inference speed. MobileNetV3-Large [32] is adopted as the backbone architecture, with an additional dropout layer to improve generalization performance, particularly when distinguishing visually similar illustrations. The model is trained with a set of cropped tag images collected during the tag registration phase. To improve robustness against variations in lighting conditions and camera angles, a variety of data augmentation techniques are applied during training. These include brightness and contrast adjustments, affine transformations (e.g., rotation and scaling), random resized cropping, and perspective warping. PoBiTag with this augmentation strategy, ensures that the model classifies tags reliably under diverse real-world usage scenarios.

However, this image-based classifier struggles to differentiate tags whose signatures primarily contain text, as text-based regions often lack distinctive visual features and tend to resemble one another, especially when they are printed in the same font, size, or layout. To address this limitation, we incorporate a text-based correction mechanism (see Figure 6). For a given segmented image, PoBiTag runs both the image-based classifier (MobileNetV3-Large) and text-based classifier, producing confidence scores c_m and c_o respectively. Specifically, for text-based classification, EasyOCR [33] is used to extract text from the segmented image and compare it with the text annotations in the training data, which are also obtained using the same OCR model during registration. The obtained texts are then encoded using character-level TF-IDF vectors incorporating bi-gram and tri-gram features and used to compute cosine similarity. Among all classes, the one whose text exhibits the highest cosine similarity to the recognized text is selected, with the corresponding similarity score employed as a confidence measure, c_o. PoBiTag then identifies a tag based on the OCR-based classification result if $c_m < 0.3$ (i.e., the image-based result is less confident) or $c_o > 0.6$ (i.e., the OCR result is deemed reliable). Otherwise, PoBiTag relies on the prediction from the image-based classifier. This hybrid approach allows the system to adapt to both image-centric and text-centric inputs, thereby improving overall classification accuracy. Upon identification, PoBiTag finally triggers the functionality associated with the recognized tag, as defined during the registration phase.

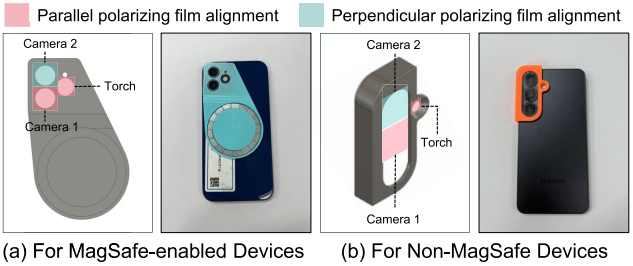


FIGURE 8. 3D-printed polarizing accessories for tag visibility enhancement.

V. IMPLEMENTATION

The prototype of PoBiTag consists of two main components: *i*) a mobile application and *ii*) a server application. The mobile application, which runs on an Apple iPhone smartphone with iOS 17.6 or later, continuously illuminates and captures surfaces containing tags using the device's rear-facing torch and camera. The captured image stream is displayed in real time on the screen, enabling users to visually detect tags embedded on the surfaces. Once a user selects a tag of interest by tapping it on the screen, the corresponding image frame and tapped location are transmitted to the server. The server application, implemented in Python, runs the proposed tag segmentation and identification techniques on a desktop computer. Upon completing the classification, the server sends the identification result back to the mobile application. Based on this result, the mobile application activates the functionality mapped to the recognized tag.

As described in Section IV-A, PoBiTag enhances the visibility of transparent tags by leveraging tags made of birefringent materials and employing polarized torch and camera. However, this reliance on specific materials and hardware limits the widespread adoption of PoBiTag in real-world environments. To address this, cellophane tape or PET stickers, everyday birefringent objects, are first utilized as tags. In addition, a mobile device's camera and torch are converted into polarized optical components by covering them with linear polarizing films. To facilitate easy attachment and removal of the films, a 3D-printed camera accessory is designed as shown in Figure 8. Specifically, the accessory securely attaches to the device using either a MagSafe [34] magnet ring (for MagSafe-enabled smartphones) or a clip-like structure (for non-MagSafe devices). It should be noted that under these configurations, PoBiTag

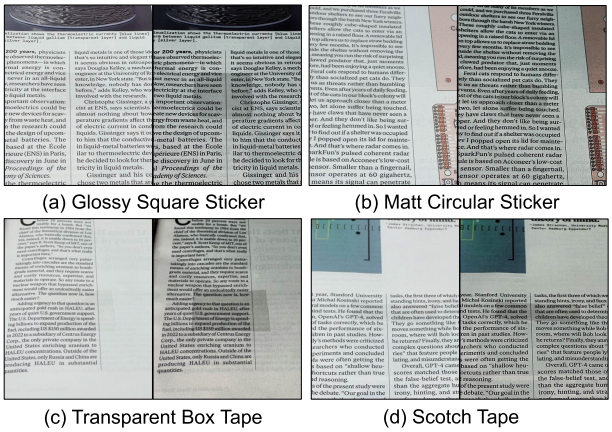


FIGURE 9. Enhanced visibility of various birefringent materials with PoBiTag. In each case, the left shows the camera view without PoBiTag, and the right shows the camera view with PoBiTag..

ensures high visibility of the tag (see Figure 9) while maintaining compatibility with existing commercial mobile devices.

VI. EVALUATION

This section evaluates the performance of PoBiTag, focusing on the following key questions:

- *Usability.* How quickly and easily can users detect transparent tags with the assistance of PoBiTag?
- *Accuracy.* How accurately can PoBiTag distinguish and identify tags?
- *Robustness.* How reliably can PoBiTag work across diverse environmental settings?

A. EVALUATION SETUP AND METHODOLOGY

During the evaluation, transparent PET stickers with a size of 30 mm × 30 mm were used as tags and attached to printed books. Specifically, seven books (B1 – B7) with different materials and contents were selected, as shown in Figure 10, each containing between 30 and 36 tags. The setup included *i*) an iPhone 12 smartphone equipped with a 3D-printed polarizing accessory (to capture the embedded tags with enhanced visibility) and *ii*) a desktop computer with an AMD Ryzen 5 5900 CPU, 64 GB RAM, and NVIDIA RTX 3090 GPU (to perform tag segmentation and identification). Note that for the tag identification, MobileNetV3-Large was trained with a set of training data collected in advance by simply capturing each tagged region from either a single viewpoint, three viewpoints, or five viewpoints (see Figure 11 for details on the viewpoints for training data collection).

Under this setup, two experiments were conducted as follows.

- *Overall performance test with real-world users.* For this test, we recruited 12 participants (5 females and 7 males with a mean age of 22.5 years). All participant recruitment and experimental procedures were carried out with the approval of the Institutional Review Board (IRB) of the institution to which the primary authors are affiliated. Participants in G1 were instructed to detect tags embedded in books B1, B3, B5, and B7 using the smartphone equipped with the polarizing accessory, while using the same smartphone without the polarizing accessory for the remaining books. The participants in G2 performed the same experiments but with the opposite configuration, using the polarized setup for books B2, B4, B6, and the non-polarized setup for the others. The entire tasks were conducted in an indoor classroom environment, where the ambient illumination was approximately 580 lux, allowing participants to freely tilt or move the smartphone to find tags. The overall performance of PoBiTag was then measured using two evaluation metrics: *tag-scanning time*, defined as the time required to detect each tag, and *identification accuracy*, defined as the proportion of correctly recognized tags after detection.
- *Robustness tests in various environments.* In this test, a single user conducted tag detection and identification but in diverse environments. By default, the user captured each tagged region in B1 three times in an indoor classroom with ambient illumination of approximately 580 lux. During data collection, the smartphone was positioned below the tagged region and tilted at an angle between 15 and 30 degrees to simulate natural handheld usage. The same procedure was then repeated under varying environmental conditions, including different smartphone orientations and lighting levels. Finally, the robustness of PoBiTag was evaluated by measuring the identification of PoBiTag for each condition.

B. OVERALL PERFORMANCE OF PoBiTag DESIGN

Figure 12 shows that PoBiTag significantly reduces tag-scanning time compared to the baseline system where a polarizing accessory for tag visibility enhancement is not applied. Participants using the baseline system often struggled to detect transparent tags, particularly in books B4 and B7, since they contain complex patterns with dark background colors. This resulted in tag-scanning time of up to 21.84 seconds (for the 97th percentile). Furthermore, none of the participants detected a tag attached to the book cover of B1, made of a highly glossy material. In contrast, all participants using PoBiTag successfully detected all embedded tags, maintaining tag-scanning time below 7.67 seconds for the 97th percentile even under visually challenging conditions. This improvement is attributed to PoBiTag's capability to provide clear visual contrast through polarization- and birefringence-based visibility enhancement, thus enabling users to easily detect and select transparent tags.

In addition to reducing tag-scanning time, PoBiTag supports precise tag identification even with minimal user effort for training data collection (see Figure 13). For

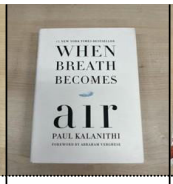
Image							
Label	B1	B2	B3	B4	B5	B6	B7
Category	Fairy tales	Fairy tales	Workbook	Magazine	Magazine	Novel	Fairy tales
Material	Thick, glossy	Thick, glossy	Thin, glossy	Thin, glossy	Thin, glossy	Thick, matt	Thick, glossy
Contents	Illustration, texts	Illustration, texts	Graphs, pictures, texts	Pictures, texts	Pictures, texts	Texts	Illustration, texts

FIGURE 10. Printed books used in experiments. We selected books with different materials and contents.

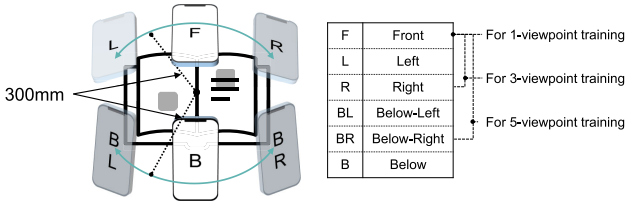


FIGURE 11. Diverse viewpoints used for data collection in experiments.

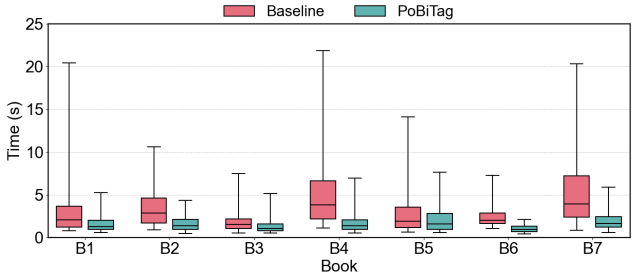


FIGURE 12. Tag-scanning time comparison. Baseline refers to PoBiTag operating without visibility enhancement (e.g., using a smartphone's camera and torch but without a polarizing accessory). The whiskers denote the 3rd and the 97th percentiles of the measured tag-scanning times.

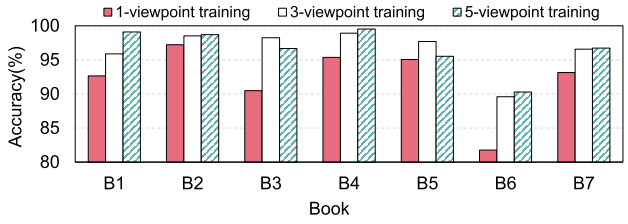


FIGURE 13. Overall tag identification accuracy of PoBiTag.

example, it achieves an average tag identification accuracy of 92.2% even when only a single training sample is collected for each tag. Augmenting training data can further improve the accuracy (e.g., 96.5% and 96.6% on average for three-viewpoint and five-viewpoint training, respectively). Notably, PoBiTag exhibited relatively lower accuracy

TABLE 2. Elapsed time of PoBiTag for tag identification. Note that it was measured on a desktop computer with an AMD Ryzen 5 5900 CPU, 64 GB RAM, and NVIDIA RTX 3090 GPU (unit = milliseconds).

	Avg.	Stdev.
Tag segmentation (w/ SlimSAM)	127.0	0.3
Text-based identification (w/ EasyOCR)	14.1	8.7
Image-based identification (w/ MobileNetV3-Large)	10.8	1.6

(81.8%) for the novel book B6 with single-viewpoint training. This is primarily because the tags in B6 contain only a small number of characters, making it difficult to extract distinctive text features for reliable identification through the text-based classifier. However, augmenting training data enables PoBiTag to correctly identify even such text-based tags as images, i.e., the image-based classifier identifies the tags with high confidence scores, thereby improving accuracy (up to 89.6% with 3-viewpoint training). Based on these results, we adopt the 3-viewpoint training configuration, as it maintains high identification accuracy while requiring less than five seconds of training data collection per tag.

1) RESPONSE TIME.

An investigation was conducted on how quickly PoBiTag can identify tags once they are selected. To this end, an average elapsed time for each processing step was measured. The processing time of PoBiTag is comparable to that of commercialized tag systems and ensures usability. As shown in Table 2, PoBiTag requires 151.9 ms to complete the tag identification process, while QR code [1], [2], [35] typically takes several tens to several hundreds of milliseconds. This result confirms that PoBiTag can support tag-based interaction with real-time responsiveness and negligible processing latency. Notably, it is expected that the processing time can be further reduced by optimizing the tag segmentation model, which is currently designed for general-purpose segmentation, and by executing the text-based and image-based classifiers in parallel.

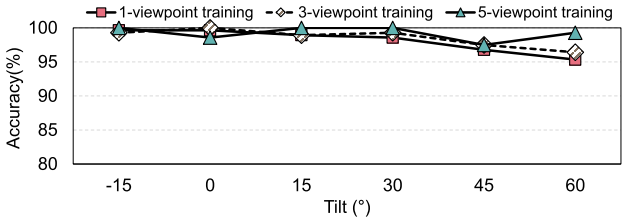


FIGURE 14. Robustness of PoBiTag against camera tilt variations.

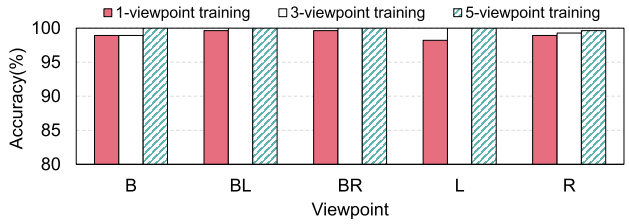


FIGURE 15. Effect of camera viewpoints in tag identification. Note that we captured tag images at each viewpoint depicted in Figure 11.

C. ROBUSTNESS OF PoBiTag DESIGN

In practice, users may operate PoBiTag under diverse environmental conditions, which can potentially affect the tag identification performance of PoBiTag. Therefore, as explained in Section VI-A, additional experiments were conducted to evaluate the robustness of PoBiTag against variations in camera orientation and lighting intensity.

1) AGAINST CAMERA ORIENTATIONS.

Variations in a camera's orientation relative to tagged areas can lead to changes in the viewing angles, which may affect the appearance of tags due to perspective distortion. To examine the robustness of PoBiTag under such conditions, its identification performance was evaluated while varying a camera's tilt and position independently.

Figure 14 illustrates that PoBiTag maintains high identification accuracy, above 95.3%, across all tested tilt angles, with only a slight decrease observed when tilts are greater than 45°. This decrease is primarily due to severe image distortion caused by steep viewing angles. However, through experiments with real-world users, it was observed that users tend to hold their smartphones with moderate tilting (e.g., between 15° and 30°) when scanning tags. In these natural cases, PoBiTag consistently achieves high accuracy even with single-viewpoint training (above 98.9%), demonstrating its practicality.

In addition, it is confirmed that PoBiTag robustly identifies tags regardless of the camera position relative to tagged areas that may affect the viewing angle. For example, as shown in Figure 15, despite the image distortions introduced by positional variations, PoBiTag maintains high identification accuracy of 98.2% even without collecting training data from multiple viewpoints. This is attributed to PoBiTag's capability to augment training data through affine transformations and perspective warping, which effectively simulate a range of viewing conditions.

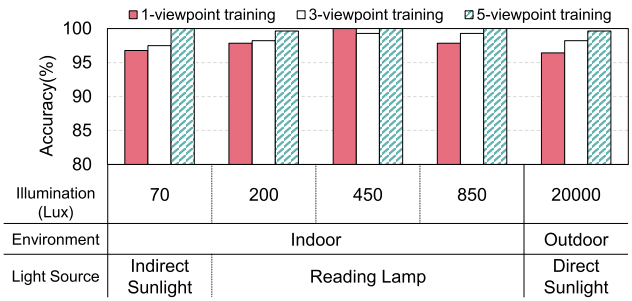


FIGURE 16. Robustness of PoBiTag against ambient lighting condition. In the indoor experiments, the ambient light intensity was controlled by adjusting the brightness level of a reading lamp placed approximately 400mm apart from tagged areas. Note that the ambient level intensity was measured using an iPhone 12 smartphone's built-in light sensor.

2) AGAINST AMBIENT LIGHTING CONDITIONS.

During tag scanning with PoBiTag, both the polarized light from the torch and ambient light (e.g., sunlight) interact with the tags, resulting in mixed reflections that are captured in the camera recordings. To evaluate the impact of ambient light on tag identification, experiments were conducted under various lighting conditions, including indoor and outdoor environments with different light source configurations as demonstrated in Figure 16. Notably, despite these varying lighting conditions, PoBiTag consistently achieved high identification accuracy (above 96.7%). This is because ambient light is typically unpolarized. As a result, ambient light passes through the polarizing filter of the camera uniformly, irrespective of whether it reflects from birefringent tags or non-birefringent backgrounds. In contrast, the polarized light emitted from the torch interacts with birefringent tags, still causing variations in polarization state that result in distinguishable intensity differences after passing through the camera's polarizer, thereby enabling robust tag detection even in the presence of ambient light.

VII. LIMITATION AND POSSIBLE MITIGATION

While PoBiTag successfully enables unobtrusive tag-based interaction through birefringence and polarization, several limitations remain, which suggest directions for future improvement.

A. AMBIGUITY ON UNIFORM BACKGROUNDS.

When tags are placed on visually uniform or empty backgrounds, PoBiTag may struggle to distinguish between different tags, as the cropped tag signatures extracted during registration may appear visually identical. One simple yet effective solution is to introduce additional visual features into tags. For instance, users may handwrite short text on transparent stickers. Alternatively, simple physical modifications to tags, such as adding small holes or cuts, can provide distinctive visual cues for classification without impacting tags' transparency or unobtrusiveness.

B. DIFFERENTIATING OVERLAPPED TAGS.

In scenarios where multiple tags overlap or are placed very close to one another, PoBiTag experiences difficulties

in accurately segmenting and classifying individual tags, as PoBiTag may detect only the outer boundary or merge multiple tags into a single region during segmentation. This problem can be mitigated by using prior knowledge of tag shapes in the tag segmentation process. For instance, we can guide PoBiTag to segment a specific shape (e.g., circles or squares) of regions. Moreover, incorporating depth sensing may provide additional spatial information to assist in resolving overlapping tag regions. This potential expansion will be explored in future work.

VIII. CONCLUSION

This study introduces PoBiTag, a novel tagging system that enables unobtrusive and customizable tag-based interaction based on the optical phenomena of polarization and birefringence. First, PoBiTag uses everyday transparent and birefringent materials, such as PET stickers or cellophane tape, as tags, allowing them to remain virtually invisible to the naked eye. However, these tags are easily detectable when captured by a mobile device's camera and torch, especially when equipped with polarizing films. The reason is that the polarized light emitted from the torch undergoes changes in its polarization state only when interacting with birefringent materials, enabling reliable tag detection. Through this visibility enhancement, PoBiTag allows users to intuitively select a tag from the captured images, accurately identify the selected tag, and activate its associated functionality. In experiments with real-world users, PoBiTag significantly reduced tag scanning time and achieved high identification accuracy across diverse environments, including varying lighting conditions, camera angles, and surface types. Future work includes further improving the practicality and scalability of PoBiTag through the design of pattern-embedded transparent tags and the development of robust multi-tag interaction capabilities.

ACKNOWLEDGMENT

The authors would like to thank development of Digital Innovative Element Technologies for Rapid Prediction of Potential Complex Disasters and Continuous Disaster Prevention.

REFERENCES

[1] T.-W. Kan, C.-H. Teng, and W.-S. Chou, "Applying QR code in augmented reality applications," in *Proc. 8th Int. Conf. Virtual Reality Continuum Appl. Ind.*, Dec. 2009, pp. 253–257.

[2] P. Kieseberg, M. Leithner, M. Mulazzani, L. Munroe, S. Schrittwieser, M. Sinha, and E. Weippl, "QR code security," in *Proc. 8th Int. Conf. Adv. Mobile Comput. Multimedia*, Nov. 2010, pp. 430–435.

[3] J. Z. Gao, L. Prakash, and R. Jagatesan, "Understanding 2D-barcode technology and applications in M-commerce-design and implementation of a 2D barcode processing solution," in *Proc. 31st Annu. Int. Comput. Softw. Appl. Conf. (COMPSAC)*, vol. 2, 2007, pp. 49–56.

[4] M. D. Dogan, A. Taka, M. Lu, Y. Zhu, A. Kumar, A. Gupta, and S. Mueller, "InfraredTags: Embedding invisible AR markers and barcodes using low-cost, infrared-based 3D printing and imaging tools," in *Proc. CHI Conf. Human Factors Comput. Syst.*, Apr. 2022, pp. 1–12.

[5] H. T. Maia, D. Li, Y. Yang, and C. Zheng, "LayerCode: Optical barcodes for 3D printed shapes," *ACM Trans. Graph.*, vol. 38, no. 4, pp. 1–14, Aug. 2019.

[6] K. D. D. Willis and A. D. Wilson, "InfraStructs: Fabricating information inside physical objects for imaging in the terahertz region," *ACM Trans. Graph.*, vol. 32, no. 4, pp. 1–10, Jul. 2013.

[7] M. Feick, X. Tang, R. Garcia-Martin, A. Luchianov, R. W. X. Huang, C. Xiao, A. Siu, and M. D. Dogan, "Imprinto: Enhancing infrared inkjet watermarking for human and machine perception," in *Proc. CHI Conf. Human Factors Comput. Syst.*, Apr. 2025, pp. 1–18.

[8] K. D. D. Willis, T. Shiratori, and M. Mahler, "HideOut: Mobile projector interaction with tangible objects and surfaces," in *Proc. 7th Int. Conf. Tangible, Embedded Embodied Interact.*, Feb. 2013, pp. 331–338.

[9] H.-J. Kim, J.-W. Kim, and T.-J. Nam, "Ministudio: Designers' tool for prototyping ubicomp space with interactive miniature," in *Proc. CHI Conf. human factors Comput. Syst.*, 2016, pp. 213–224.

[10] D. K. Rosner and K. Ryokai, "Spyn: Augmenting the creative and communicative potential of craft," in *Proc. SIGCHI Conf. Human Factors Comput. Syst.*, Apr. 2010, pp. 2407–2416.

[11] Z. Chen, G. Fang, X. Ma, and X. Wang, "SlimSAM: 0.1% data makes segment anything slim," 2023, *arXiv:2312.05284*.

[12] M. Fiala, "ARTag, a fiducial marker system using digital techniques," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, vol. 2, Jun. 2005, pp. 590–596.

[13] E. Olson, "AprilTag: A robust and flexible visual fiducial system," in *Proc. IEEE Int. Conf. Robot. Autom.*, May 2011, pp. 3400–3407.

[14] B. Atcheson, F. Heide, and W. Heidrich, "CALTag: High precision fiducial markers for camera calibration," in *Proc. VMV*, vol. 10, Jan. 2010, pp. 41–48.

[15] Z. Wang, S. Qiu, Q. Chen, N. Trayan, A. Ringlein, J. Dorsey, and H. Rushmeier, "AniCode: Authoring coded artifacts for network-free personalized animations," *Vis. Comput.*, vol. 35, nos. 6–8, pp. 885–897, Jun. 2019.

[16] M. Maida, A. Crescini, M. Perronet, and E. Camuffo, "Claycode: Stylable and deformable 2D scannable codes," 2025, *arXiv:2505.08666*.

[17] Z. Tian, C. J. Carver, Q. Shao, M. Roznere, A. Q. Li, and X. Zhou, "PolarTag: Invisible data with light polarization," in *Proc. 21st Int. Workshop Mobile Comput. Syst. Appl.*, Mar. 2020, pp. 74–79.

[18] D. Li, A. S. Nair, S. K. Nayar, and C. Zheng, "AirCode: Unobtrusive physical tags for digital fabrication," in *Proc. 30th Annu. ACM Symp. User Interface Softw. Technol.*, Oct. 2017, pp. 449–460.

[19] M. D. Dogan, F. Faruqi, A. D. Churchill, K. Friedman, L. Cheng, S. Subramanian, and S. Mueller, "G-ID: Identifying 3D prints using slicing parameters," in *Proc. CHI Conf. Human Factors Comput. Syst.*, Apr. 2020, pp. 1–13.

[20] C. Getschmann and F. Echtler, "Seedmarkers: Embeddable markers for physical objects," in *Proc. 15th Int. Conf. Tangible, Embedded, Embodied Interact.*, Feb. 2021, pp. 1–11.

[21] D. Li, D. I. W. Levin, W. Matusik, and C. Zheng, "Acoustic voxels: Computational optimization of modular acoustic filters," *ACM Trans. Graph.*, vol. 35, no. 4, pp. 1–12, Jul. 2016.

[22] L. He, G. Laput, E. Brockmeyer, and J. E. Froehlich, "SqueezaPulse: Adding interactive input to fabricated objects using corrugated tubes and air pulses," in *Proc. 11th Int. Conf. Tangible, Embedded, Embodied Interact.*, Mar. 2017, pp. 341–350.

[23] C. E. G. Tejada, O. Fujimoto, Z. Li, and D. Ashbrook, "Blowhole: Blowing-activated tags for interactive 3D-printed models," in *Proc. Graph. Interface*, Jun. 2018, pp. 131–137.

[24] C. Harrison, R. Xiao, and S. Hudson, "Acoustic barcodes: Passive, durable and inexpensive notched identification tags," in *Proc. 25th Annu. ACM Symp. User Interface Softw. Technol.*, Oct. 2012, pp. 563–568.

[25] G. Ulian and G. Valdrè, "Study of the variation of the optical properties of calcite with applied stress, useful for specific rock and material mechanics," *Sci. Rep.*, vol. 12, no. 1, p. 299, Jan. 2022, doi: 10.1038/s41598-021-04471-6.

[26] A. D. Slepkov, "Quantitative measurement of birefringence in transparent films across the visible spectrum," *Amer. J. Phys.*, vol. 90, no. 8, pp. 625–634, Aug. 2022, doi: 10.1119/5.0087798.

[27] S. M. Yang, S. Hong, and S. Y. Kim, "Optical, mechanical, and photoelastic anisotropy of biaxially stretched polyethylene terephthalate films studied using transmission ellipsometer equipped with strain camera and stress gauge," *J. Polym. Sci. Part B, Polym. Phys.*, vol. 57, no. 3, pp. 152–160, Feb. 2019.

[28] I. Min and K. Yoon, "An experimental study on the effects of injection-molding types for the birefringence distribution in polycarbonate discs," *Korea-Australia Rheology J.*, vol. 23, no. 3, pp. 155–162, Sep. 2011.
[29] A. M. Reza, "Realization of the contrast limited adaptive histogram equalization (CLAHE) for real-time image enhancement," *J. VLSI signal Process. Syst. Signal, Image Video Technol.*, vol. 38, no. 1, pp. 35–44, Aug. 2004.
[30] A. Kirillov, E. Mintun, N. Ravi, H. Mao, C. Rolland, L. Gustafson, T. Xiao, S. Whitehead, A. C. Berg, W.-Y. Lo, P. Dollár, and R. Girshick, "Segment anything," 2023, *arXiv:2304.02643*.
[31] G. Fang, X. Ma, M. Song, M. Bi Mi, and X. Wang, "DepGraph: Towards any structural pruning," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2023, pp. 16091–16101.
[32] A. Howard, M. Sandler, B. Chen, W. Wang, L.-C. Chen, M. Tan, G. Chu, V. Vasudevan, Y. Zhu, R. Pang, H. Adam, and Q. V. Le, "Searching for MobileNetV3," *CoRR*, pp. 1314–1324, Oct. 2019.
[33] JaidedAI. (2020). *Easyocr*. Accessed: Jun. 11, 2025. [Online]. Available: https://github.com/JaidedAI/EasyOCR
[34] Apple Inc. (2020). *MagSafe*. Accessed: Jun. 11, 2025. [Online]. Available: https://www.apple.com/magsafe/
[35] D. R. Corpuz, L. G. Rosario, J. P. C. Cempron, P. L. Lozano, and J. Ilao, "QR code detection with perspective correction and decoding in real-world conditions using deep learning and enhanced image processing," in *Proc. Copyright*, Jan. 2025, pp. 685–690.



SEOHYEON PARK received the B.S. degree from the Department of Computer Science and Engineering, Chung-Ang University, Seoul, South Korea, in 2023. She is currently pursuing the M.S. degree with the School of Computer Science and Engineering, Chung-Ang University. Her research interests include cyber-physical systems, signal processing, and mobile and ubiquitous systems.



SEUNGHUN CHAE received the B.S. degree from the Department of Computer Science and Engineering, Chung-Ang University, Seoul, South Korea, in 2023. He is currently pursuing the M.S. degree with the School of Computer Science and Engineering, Chung-Ang University. His research interests include human-computer interaction, near-field communication using sound, and ubiquitous systems.



JAEMIN CHOI received the B.S. and M.S. degrees from the Department of Computer Science and Engineering, Chung-Ang University, in 2023 and 2025, respectively. His research interests include cyber-physical systems, signal processing, and mobile and ubiquitous systems.



HYOSU KIM (Member, IEEE) received the B.S. degree from the Department of Computer Engineering, Sungkyunkwan University, in 2010, and the M.S. and Ph.D. degrees from the School of Computing, Korean Advanced Institute of Science and Technology (KAIST), Republic of Korea, in 2012 and 2018, respectively. He is currently an Associate Professor with the School of Computer Science and Engineering, Chung-Ang University, Seoul, South Korea. His research interests include human-computer interaction, mobile and ubiquitous computing, and cyber-physical sensing.