AccessLens: Auto-detecting Inaccessibility of Everyday Objects

Nahyun Kwon
nahyunkwon@tamu.edu
Texas A&M University
College Station, Texas, USA

Qian Lu
qianlu@tamu.edu
Texas A&M University
College Station, Texas, USA

Muhammad Hasham Qazi
qazihasham@gmail.com
Texas A&M University
College Station, Texas, USA

Joanne Liu
joannechanliu@tamu.edu
Texas A&M University
College Station, Texas, USA

Changhoon Oh
changhoonoh@yonsei.ac.kr
Yonsei University
Seoul, Korea

Shu Kong
shu@tamu.edu,skong@um.edu.mo
Texas A&M University
University of Macau, United States

Jeeeun Kim
jeeeun.kim@tamu.edu
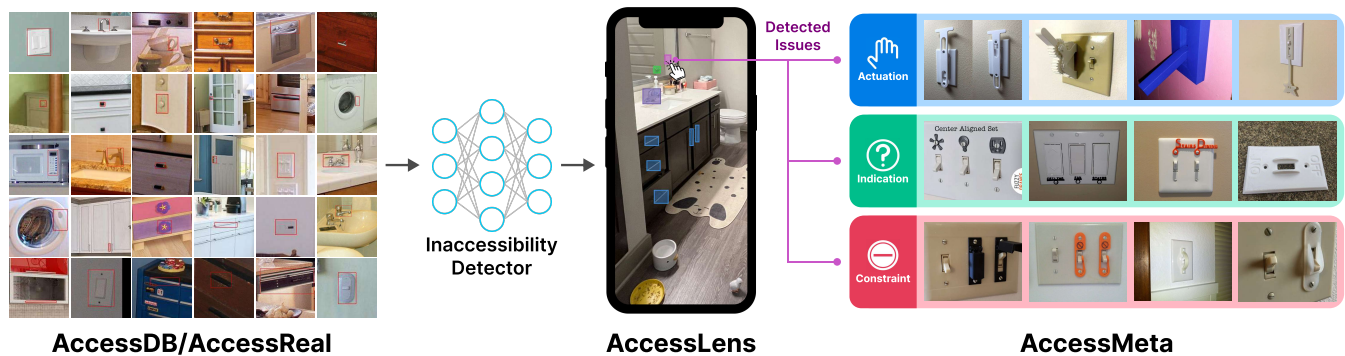Texas A&M University
College Station, Texas, USA

Figure 1: AccessLens system overview. AccessLens provides a mobile toolkit to scan indoor scenes and detect inaccessibility in everyday objects. Inaccessibility detection is developed on our dataset AccessDB and AccessReal, consisting of indoor scene images annotated with inaccessibility classes on daily objects. We contribute AccessMeta, a metadata that categorizes 3D assistive designs, enabling auto-suggestions to improve daily accessibility.

ABSTRACT

In our increasingly diverse society, everyday physical interfaces often present barriers, impacting individuals across various contexts. This oversight, from small cabinet knobs to identical wall switches that can pose different contextual challenges, highlights an imperative need for solutions. Leveraging low-cost 3D-printed augmentations such as knob magnifiers and tactile labels seems promising, yet the process of discovering *unrecognized* barriers remains challenging because disability is context-dependent. We introduce AccessLens, an end-to-end system designed to identify inaccessible interfaces in daily objects, and recommend 3D-printable augmentations for accessibility enhancement. Our approach involves training a detector using the novel AccessDB dataset designed to automatically recognize 21 distinct Inaccessibility Classes (e.g., bar-small and round-rotate) within 6 common object categories (e.g., handle and knob). AccessMeta serves as a robust way to build a comprehensive dictionary linking these accessibility classes to open-source 3D augmentation designs. Experiments demonstrate our detector's performance in detecting inaccessible objects.

CCS CONCEPTS

• **Human-centered computing** → **Collaborative and social computing systems and tools**; **Interactive systems and tools**; • **Computing methodologies** → *Computer vision*.

KEYWORDS

3D assistive design, object detection, end-user interface

1 INTRODUCTION

While the traditional definition of disability has revolved around individuals' varied abilities, understanding disability as *'mismatched interactions'* [48] emphasizes diverse contexts that can create barriers within environments. Consider someone with a wrist injury struggling with everyday tasks like opening a water bottle or using a toothbrush single-handedly; new parents suddenly recognize potential hazards at home, such as electric outlets. However, recognizing such contextual disability and proactively rectifying them remains challenging for inexperienced users because they use prior experiences that could be biased. It is non-trivial to foresee unfamiliar interaction scenarios (e.g., managing everyday tasks by being one-handed), leading them to cope with difficulties without promptly addressing interaction challenges.

"If the design is accessible, everyone benefits" [1]; the accessibility community has highlighted the importance of engaging everyone in improving accessibility. Traditional approaches to raising awareness and fostering proactive efforts focused on cultivating empathy and mutual understanding among non-disabled individuals. The goal was to evoke recognition of unnoticed discomfort inherent in daily interfaces, particularly from the perspective of individuals with disabilities [19, 46, 57, 58, 60]. However, these approaches had inherent limitations in simulating disabilities, which could inadvertently lead to biases and cognitive gaps against individuals without disabilities [53]. Although well-structured textual guidelines and compliances [16, 28, 56] encompass exhaustive domain knowledge from experts, those remain static, exclaiming the need for interactive systems. However, while the disability is context-dependent, implying that anyone can experience challenges without permanent disability, the latest AI-powered interactive tools [55, 68] predominantly focus on specific target groups, such as wheelchair users or older adults, missing the contextual variances, i.e., temporary and situational cases [48]. Moreover, many solutions entail renovation or replacements, which is often costly thus mentally burdening, limiting the practicality/applicability of existing tools in promoting pro-social behaviors. There remain three major user challenges:

- Which objects are inaccessible?
- Why and when do these objects become inaccessible?
- How can a user without prior experiences identify them and find appropriate solutions?

We introduce **AccessLens**, an end-to-end system to automate detecting contextual barriers from everyday objects, and suggest 3D-printed assistive augmentations. Figure 1 shows system overview. AccessLens is built upon novel datasets, **AccessDB/AccessReal** to train inaccessibility detectors, and **AccessMeta**, metadata to understand interaction types and required human capabilities of physical objects presented as their interaction attributes. As existing datasets (e.g., [26, 38, 77]) with indoor scene images do not articulate inaccessibility to automate detection, AccessDB was built to imbue accessibility knowledge using 21 Inaccessibility Classes (IC). Designed to foster understanding of how 3D assistive augmentations can resolve contextual disabilities, AccessMeta provides the link between 3D augmentations and interaction types/contexts of existing objects, such as a lever extension for a door knob that removes sophisticated motor skills (Figure 2a-b) and an arm-pull extension for a lever for an alternative operation (Figure 2c-d).
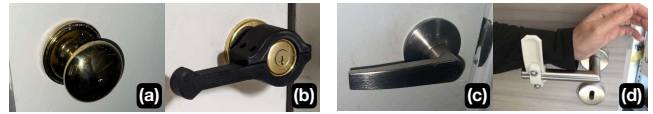

Figure 2: (a) A round knob's accessibility can be improved by (b) lever extension [71] while (c) a lever handle's accessibility is improved by an (d) arm extension [30]. Everyday objects portray different accessibility barriers to people under different contexts.

In sum, our contributions are three-fold:

- A holistic survey of large-scale 3D assistive augmentations in online repositories and understanding of their interaction properties, resulted in **AccessMeta**, a metadata to autoclassify them;
- **AccessDB & AccessReal**: A dataset for auto-detection of inaccessible objects and parts from indoor scenes with 10k annotated objects under 21 Inaccessibility Classes with realistic high-res dataset for testing;
- **AccessLens**: End-user system to detect inaccessibility and to obtain design recommendations through 3D printed augmentations to update legacy objects

We evaluate our contributions through user studies and technical experiments. First, a preliminary user evaluation of the AccessLens system prototype helps understand how AccessLens enhances awareness and willingness to take pro-social behaviors. Second, we assess an end-to-end pipeline—capturing the indoor environment to retrofitting 3D augmentations—with inexperienced users and two experts in assistive technology. The evaluation of AccessMeta engaged crowdworkers in annotating the dictionary with 280 3D augmentations. We also evaluate AccessDB/AccessReal with off-the-shelf detectors.

Our vision for AccessLens is to empower users with limited awareness to recognize hidden daily accessibility challenges thus to be more attentive to daily challenges under diverse contexts and extents. AccessLens does not require diagnosed disability, prior experience, and domain expertise to recognize inaccessibility. Figure 3 shows our scope on target demographics compared to existing approaches.
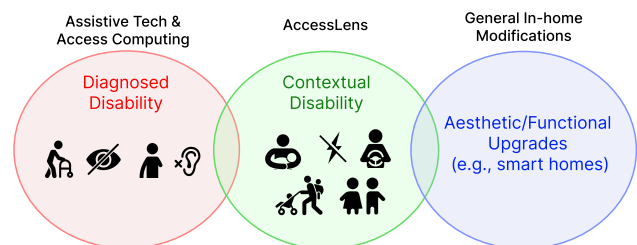

Figure 3: AccessLens's target user scope compared to existing assistive technique works and general in-home modifications. AccessLens supports users with limited awareness but who can easily become disabled under various contexts.

	Compliances/ guidances e.g., [16, 56]	MS Inclusive Guidebook [48]	Project Sidewalk [64]	Homefit AR [55]	RASSAR [68]	AccessLens (ours)
Interactive	✗	✗	✓	✓	✓	✓
Indoor accessibility	✓	▲	✗	✓	✓	✓
Contextual disability	✗	✓	✗	✗	✓	✓
Auto detection	✗	✗	✓	✓	✓	✓
Interaction type detection	✗	✗	✗	▲	✗	✓
Low-cost adapations	✗	✗	✗	✗	▲	✓

Table 1: Position of AccessLens compared against prior works.

2 RELATED WORK

2.1 Interactive Tools to Evaluate Accessibility

There exist numerous standards and normative tools to help non-experts learn cumulative knowledge. The Americans with Disabilities Act (ADA) Standards for Accessible Designs [16] and the International Building Code (IBC) [28] represent comprehensive frameworks to alleviate mobility challenges. Increasing interests are in their interactivity, for instance, improving indoor access for older adults through interactive systems [11, 18, 42, 43, 55]. Homefit AR [55, 56] guides users through questionnaires to precisely locate issues with object types and recommends alternatives with better access. The closest prior work of ours is RASSAR [68], a mobile AR application to assess indoor accessibility upon standards such as low tables, narrow entryways, and dangerous items exposed. While these works respond to the needs of special interest groups (e.g., older adults and wheelchair users), a broader population is often excluded, since they have not experienced disabilities and could overlook contextual or situational disabilities. We are to provoke solutions with an emphasis on the engagement of a more diverse community in creating accessible and accommodating indoor spaces.

2.2 Advancing Accessibility: Beyond Empathy and Simulations

Fostering empathy is discussed in many disability studies to elevate awareness about the lived experiences of disabled people [8, 57, 58]. While simulating disabilities such as blindfolding [60], having color-blind effects [19], or trying wheelchairs [46] has gained popularity, disability advocates disparage simulated disability [8, 27]; as it is difficult to accurately replicate the real experiences [4]. Empathy alone may not suffice to sustain attention [27], simulations may inadvertently create biases or distress [53], resulting in perpetuated ableism [25]. More recent focus is on co-designing *with* people with disabilities (e.g., [7, 32, 33, 74]); e.g., citizens, healthcare professionals, and designers co-design personalized healthcare solutions [33], sighted and blind participants design building navigation together [32]. Collective efforts to enhance the user experience can extend the impact beyond individuals with disabilities alone, encompassing different abilities of all [48, 61, 63]. Unfortunately, there have been only little to no systems to engage non-experienced users to cultivate inclusivity. Our approach is motivated by "*design for one, expand to all*" and "*learning from adaptations*" [48] to promote awareness through noticing and better designs.

2.3 Indoor Scene Understanding

Visual perception of indoor places is a critical first step to improving one's quality of life [65]. Various datasets were released to train detectors. Some early datasets such as MIT indoor scenes [59] and SUN RGB-D [66] have advanced techniques to train recognition models. Synthetic datasets such as HyperSim [62] can further advance recognition models with their variety and quantity. While there exist relevant datasets such as Gibson [76], offering a virtual visual navigation platform, PartNet [51] with focus on part of indoor objects, and BEHAVIOR-1K [34], data for embodied AI systems to foster human-robot interaction, none have centered interaction types to assess their accessibility and user contexts. We find ADE20K [77], which is a large-scale indoor scene dataset with hierarchical annotations of objects in images at the pixel level, promising. Refining the hierarchical taxonomy of objects and parts by ADE20K includes object categories and parts, we curate datasets by re-annotating potentially inaccessible objects to train and evaluate inaccessibility detectors.

2.4 3D-Printed Augmentations to Improving Access to Legacy Objects

While it is not feasible to replace *all* existing objects overnight [3, 35], 3D-printed assistive designs [5, 10] promises low-cost, custom solutions to redress everyday interaction challenges (e.g., [10, 14, 22]). These adaptations can range from magnifying cabinet knobs for improved grip (e.g., 'ThisAbles' [67]) to self-serving medicine dispensers [3]. Similar to the modular approach employed in the modern software engineering paradigm, wherein updates are selectively applied only where changes are necessary [54], the augmentation allows for unit-by-unit enhancements tailored to specific needs. Barriers to 3D printing have been significantly lowered [9], existing works studied motivations behind online communities sharing assistive 3D designs [10] and proposed computational customization solutions (e.g., [14]). While documents based on similarity can classify shared designs' objectives [36], current search relies on designer-created descriptions, often failing users to explore viable designs to rectify hidden inaccessibility that is not obvious to those without diagnosed disabilities. Discovering suitable designs heavily relies on keyword-based searches, relying on the textual information provided by the authors: titles, descriptions, and tags only. We propose novel metadata to categorize existing 3D assistive augmentations for better identification of solutions.

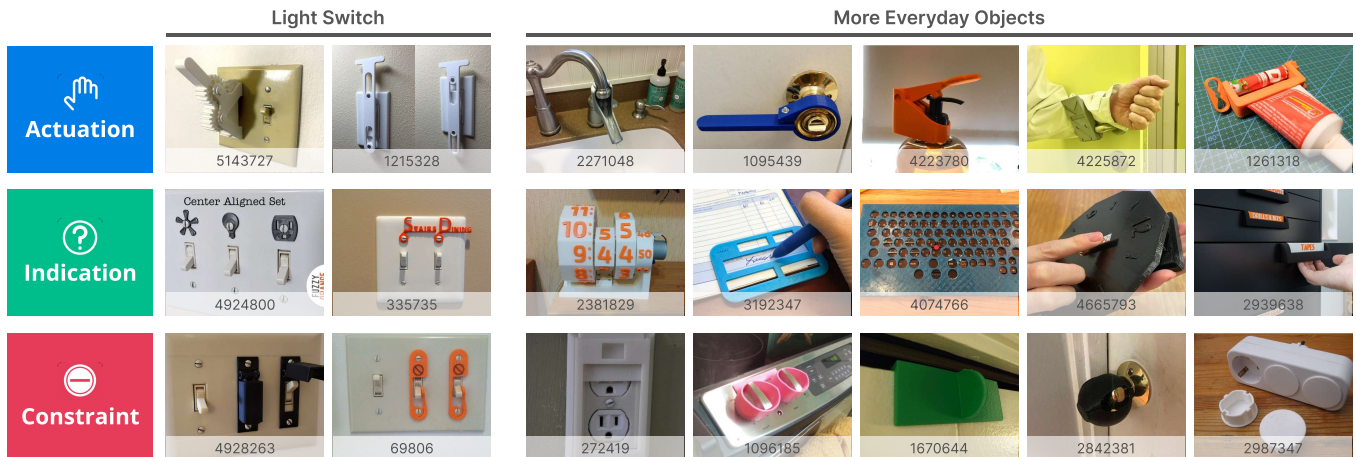In sum, Table 1 summarizes the position of Accesslens.

Figure 4: Examples of 3D assistive augmentations that belong to three categories, obtained from our in-the-wild survey with iterative affinity diagramming. Each design has a *thing_id* at the bottom, and the design page can be located at https://www.thingiverse.com/thing:*thing_id*. Examples show that various challenges, such as motor and sensory barriers, can present even for one object. 3D augmentations are actively used to address challenges without requiring total replacement.

3 DESIGNING ACCESSLENS

We introduce two design studies; the first investigates how people currently make changes to their environments by adopting 3D printed augmentations through an in-the-wild survey. Taking account into the design objectives and interactions that entail, the second examines our design probe with 8 participants whether the system helps naive users interpret daily accessibility challenges differently.

3.1 Design Study #1. Understanding Interaction Contexts: In-the-Wild Survey

We conducted an exploratory survey on Thingiverse [29], to gain insights into why individuals are motivated to create 3D assistive augmentations and modify existing physical objects to address specific contextual or situational interaction challenges. First, we listed several indoor objects that are very common around us, including door knobs, light switches, etc. Then we retrieved 3D designs that are for those objects, indicating 3D designs tend to augment targeted real-world objects from Thingiverse. We employed an iterative process of affinity diagramming, which was collaboratively performed by four of our authors. In the affinity diagramming process, we classified augmentations considering three primary criteria: (1) their intended objective, which refers to the barriers the augmentations aim to address, (2) the type of objects the augmentations target, and (3) any related motions or actions associated with their use. Our empirical findings revealed that even for objects that are under the same class (e.g., door knob/handle, light switch), the augmentations are much more diverse due to differences in the object's type (e.g., single toggle light switch vs. rocker switch). This diversity emanates from shapes, motions, and objectives, which inspired us to develop AccessDB, our refined dataset with inaccessibility classes of AccessMeta. This iterative affinity study resulted in three high-level functions of adaptations as follows and example augmentations are shown in Figure 4.

- **Reducing motor requirements, change needed motion types [Actuation]:** Designs that shift types of motions needed to operate (e.g., rotation to linear push) or reduce workload (e.g., reduce required power to manipulate interfaces, or allow one hand instead of two hands); for people with motor limitations.
- **Furnishing with visual/tactile cues [Indication]:** Designs that create multi-modal functions for identification, providing labels (e.g., switch identifiers, toggling sound); for people with sensory limitations.
- **Adding constraints [Constraint]:** Designs that prevent a targeted population from operating a task by limiting their operation mainly due to safety reasons (e.g., cabinet lock, switch lock, stove knob stopper); for people with cognitive limitations or child-access/child-proof products.

3.2 Design Study #2. Design Probe

We developed the prototype of the AccessLens and conducted a comparative study to assess its validity and advanced features over the baseline, MS Inclusive Design Guidebook [48]. Compared to other normative tools that are targeted to diagnosed disabilities, e.g., ADA Standards for Accessible Design [16], the MS guidebook is the foremost design guideline that argues accessibility as a universal daily challenge *for all*, encouraging recognizing exclusion, extending the inaccessibility concept to contextual from a permanent problem. Herein, the disability is discussed not as a personal health condition, but as *'mismatched human interaction'* which we see the potential to rectify through augmentations. Thinking of solutions for those situational disabilities can allude to a design for one that can benefit all [61], empowering people to learn from diversity. AccessLens prototype (Figure 5) includes objects with detected potential accessibility challenges. Tapping on objects, the system displays relevant 3D augmentations depending on contextual needs. We provide the contexts through a catalog approach, helping users learn from viewing adaptations list, which also presents design implications to nurture people's understanding of solutions.
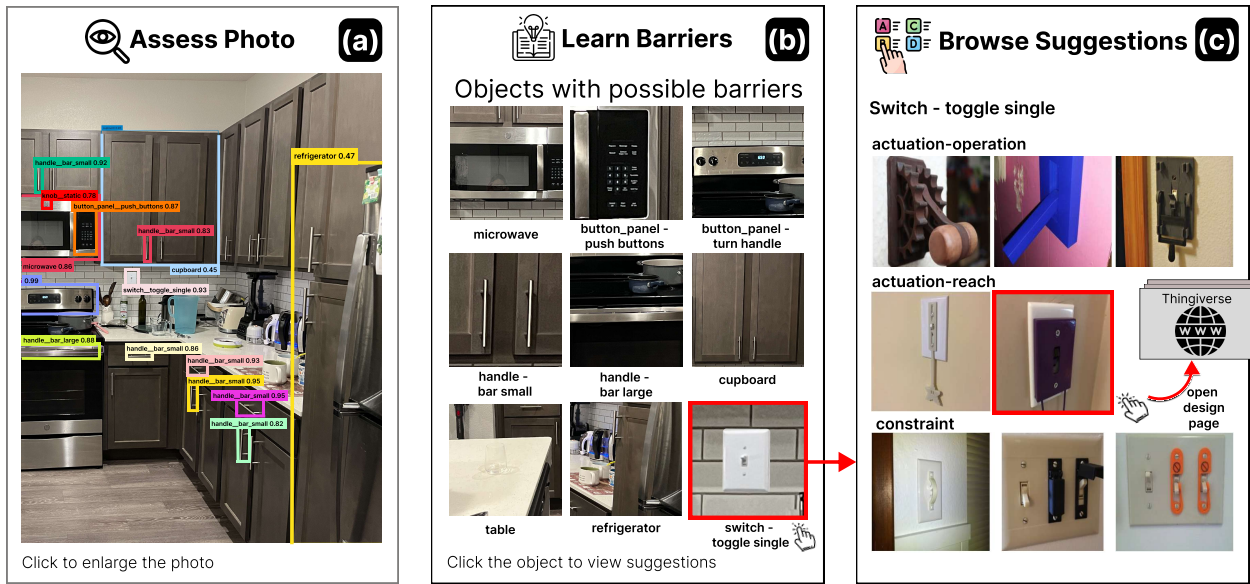
Figure 5: AccessLens prototype overview. AccessLens allows users to scan an uploaded photo (a), view the detected inaccessible objects (b), and upon a click of a detected object, browse through the available suggestions (c).

3.2.1 Participants. We recruited 8 participants from various backgrounds, including researchers who are not in the accessibility domain (*N*=5), educators (middle/high school teacher, college professor, *N*=3). Two self-identified as older adults (*N*=2). Aligning with our target users who do have limited experiences in the accessibility concepts, we recruited participants without diagnosed disabilities nor knowledge of accessibility study. We observed whether AccessLens promotes *"thinking about daily inaccessibility".*

3.2.2 Procedure. We chose a within-subject study. We counterbalanced the conditions to reduce learning effects; half of the participants started with the baseline condition, and the other half started with the experimental condition. The study sessions began with a pre-task interview. Participants then completed the same tasks under two conditions and finally, took a closing interview. In the pre-task interview, participants shared their prior experiences when they encountered difficulties in interacting with everyday objects or witnessed someone else having issues. They were also asked if they had implemented any solutions to address such barriers. One study condition is the Baseline condition, where participants access the link to the introduction video for MS Inclusive Design [49] and the MS Inclusive 101 guidebook [48] (MS guidebook, hereinafter). Participants were allowed to spend enough time reading the guidebook, without any time restrictions. During the task, participants were presented with indoor scene images and identified the objects that could present potential accessibility barriers. They were then asked to propose solutions. Subsequently, participants were asked to rate each suggestion on a 5-point Likert scale. Participants were encouraged to use any necessary online resources (e.g., YouTube and Google Search) in the baseline. In the experimental condition, participants used AccessLens but were not permitted to access other online resources. As shown in Figure 5, the AccessLens displays indoor scene images of chosen, highlighted objects that could be inaccessible and offers applicable solutions. The task was repeated

with a different indoor scene image. After both conditions, a brief interview followed for 15 minutes. We investigated their perceived usefulness by providing a survey questionnaire measuring three sub-metrics on a 5-Likert scale: (1) recognizing inaccessible objects, (2) comprehending related contexts, and (3) identifying solutions. All responses and comments were documented for analysis. The entire session took 1.5 hours on average, not exceeding 2 hours. The study has been approved by the institution's review board (IRB No.: IRB2023-0648)
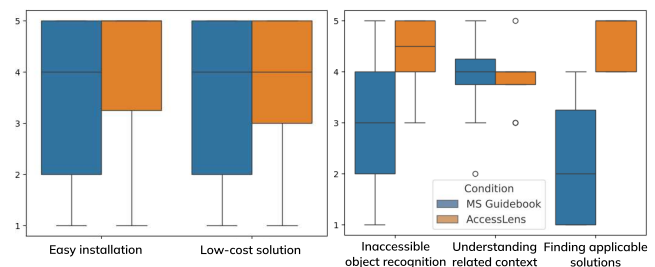
Figure 6: (a) 3D augmentation recommendations are assessed by two sub-metrics of easy installation and low-cost solution. (b) Perceived helpfulness is assessed by inaccessible object recognition, understanding contexts, and solution retrieval.

3.2.3 Findings & Implications.
#1. Ableism: Overlooked Inaccessibility and Gaps between noticing and an Action. P2 shared the story of their mother suffering from an ankle injury, leading her to stay seated at home until recovered. All often relied on family members for assistance, such as getting dressed with the help of a sibling (P2), and tried to circumvent challenges by struggling to use a non-dominant hand

(P5) which was not perceived as a 'disability' at that moment. Internalized ableism might explain this, where individuals may think disability *"has to cross some threshold of difficulty or suffering to count"* [25] and do not think of their constraints as living disabilities to be addressed with solutions. Standing out to those who do not present diagnosed disabilities, ableism eventually misses the opportunities to renovate their environment for future contexts.

#2. Learning from Adaptations. Participants noted that design recommendations make them infer contexts albeit no explicit descriptions were provided. Several participants liked the persona spectrum presented in the MS guidebook, how different disabilities can relate to each other, broadening their understanding of disability. P5 mentioned that he now recalls he was temporarily disabled. AccessLens achieved the same effect by cataloging various augmentations. It encouraged participants to "learn from adaptations" [48]. Many were surprised by the variety of AccessLens recommendations, admitting they had not considered accessibility issues those designs could negate. *"I hadn't thought these [could be an] issue before I saw these designs"* (P2, P4). P3 felt gratitude for the detection & suggestion together. *"When I only saw the photo of the room [...], even when I see the detected objects, I didn't know which contexts it can pose barriers. When I saw the suggestions, I could imagine in which situations it can be helpful and what the objective is [of those or similar designs]"*. We confirmed that presenting *better designs* can inspire and let users comprehend the diversity. P5 preferred AccessLens highlighting its transformative impacts; *"We usually think only of the disabled [when we were asked to think about disability]. AccessLens makes me think that even the non-disabled can get help and apply the solutions in their environments"*.

#3. Mental load in Disability Accommodation. We questioned the estimated installation expenditure under two sub-metrics: *(1) easy installation* and *(2) low-cost solution*. Figure 6a shows their estimation in easiness/affordability. The average score does not indicate notable differences, possibly due to the learning effect; participants who experienced the AccessLens first tended to use their knowledge obtained during the following baseline condition. Participants who began with the baseline guessed replacement or extensive renovations as the sole solutions. While perceived difficulty and cost varied among participants, they projected high cost and effort for replacements. Most participants were curious about market products in the baseline condition. P5 imagined aggregated dials with small labels on a kitchen stove could be confusing for older adults, thinking individual knobs for each burner would be helpful, but questioning whether he could get one off-the-shelf. In contrast, all found AcceeLens recommended 3D augmentations straightforward and cost-effective. *"I thought that we always needed complete replacements or renovations [...] Reviewing the suggestions, I realized that these solutions can be easily installed so I really want to install them, [e.g., childproof augmentations] to ensure safety"* (P3).

On the other hand, in baseline, none utilized external sources not knowing what and how to search, implying low engagement. Only P1 tried general search keywords (e.g., assistive bathroom, accessible bathroom). *"I had to brainstorm to find the solutions [on my own]. Even with online resources allowed, I believe it wouldn't be that helpful because I need to know what to search for"* (P1). This signifies the higher mental load keeps users from engaging in solution-seeking/adaptations.

#4. Written Guidebook vs. Interactive System. In the closing interview, participants evaluated two conditions across three sub-metrics: (1) the ability to recognize inaccessible objects, (2) understanding related contexts with barriers, and (3) retrieving applicable solutions. Figure 6b summarizes participants' assessment, showing AccessLens outperforms the guidebook in terms of detecting inaccessible objects and seeking solutions.

#5. Interaction Design. In addition, participants desired, (1) implementing on a mobile reduces the user experience gap between capturing photos and inspection. (2) Context-based filtering to reconcile accessibility evaluations to certain scenarios, increasing the system's versatility. (3) A summary view of all detection with bounding boxes to simplify the inspection process for a quick overview at a glance, and also to grasp the objectives of proposed accessibility enhancements swiftly. (4) Supplementary explanations for the categorization, i.e., AccessMeta categories, will enhance in-depth appreciation of the suggestions and their design intention. (5) A tutorial or instructional guide on how to capture photos would help users in providing clear and relevant images. These collective enhancements were reflected in AccessLens improvements. We elaborate on an improved design as in Figure 7.

3.3 Design & Implementation Considerations

Consideration #1: One-shot Image Input From the HCI perspective, allowing users to upload a single photo of an indoor scene would offer a more pleasant experience, considering that our target users might not know where to focus. While detection performance can benefit from multiple photos of the indoor scene, it is more friendly for users to take a single photo of the entire room or scanner view to check whether there exist any inaccessibility concerns. We target one-shot imagery of indoor scenes of interest as input, i.e., a panoramic scan of a bathroom, living room, and office space.

Consideration #2. Semantic Understanding of Parts To assist users with different needs, detecting part (doorknob from a door) and discerning the *type* of the object (doorknob vs. lever) is critical to articulate contextual barriers beyond simple object detection. The system must detect target objects and the parts where actual user interaction occurs, since each presents unique barriers with associated interaction types, for example, a knob for grab-pull vs. a knob for grab-rotate. Therefore, the image dataset must contain indoor scenes with part-level annotations.

Consideration #3: Recognizing Disability Attributes. Various contexts change the way that people with a wide spectrum of capabilities interact with everyday objects; for a graphic designer wearing a splint due to chronic wrist pain, a door knob is not accessible as it requires hard grasping to rotate. People are often frustrated with a panel with identical toggle switches; without labels, they are forced to recall targets or try to get the right one turned on, sometimes causing safety breaches. The disability context attributes of the objects might fortify the existing dataset. In sum,

(1) A user should be able to use a general view of scenes as input instead of a focused view of interested objects.
(2) The system must be able to semantically understand the detected objects (e.g., cabinet knob vs. door knob).
(3) A new dataset must account for understanding various inaccessibility contexts beyond object/instance detection.
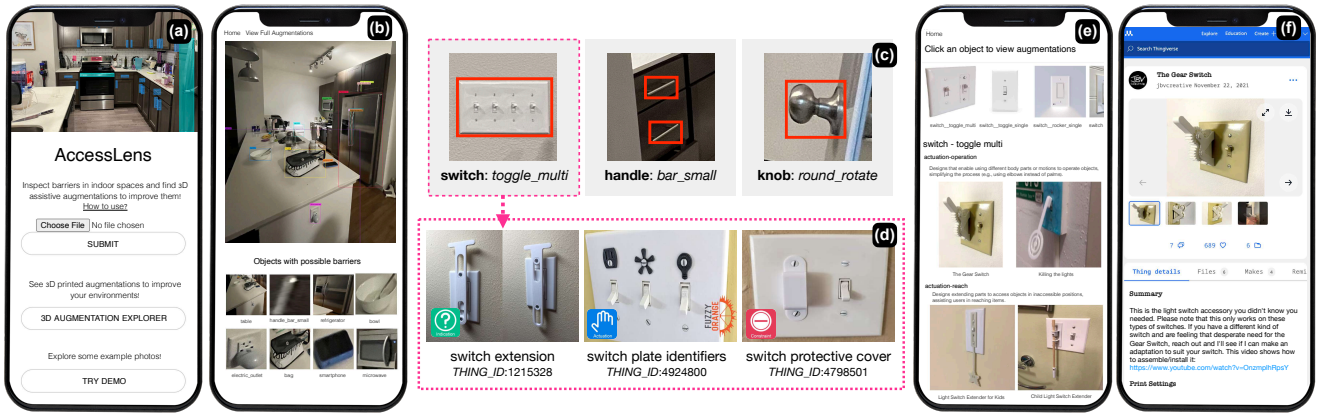
Figure 7: AccessLens: (a) main page, (b) an indoor image with detected objects with barriers, (c) example inaccessibility classes, (d) 3D-printed augmentations classified with AccessMeta, (e) a 3D augmentation explorer to view full suggestions, and (f) redirecting to the design page for details.

4 ACCESSLENS

AccessLens comprises AccessDB, AccessMeta, and the end-user toolkit, designed to seamlessly work together to assist end users in addressing accessibility challenges. The center of its functionality is AccessDB, a dataset used to train the inaccessibility detector, which analyzes images captured by users via a mobile user interface 1. The detector identifies inaccessible objects within various possible contexts. Leveraging AccessMeta, AccessLens suggests the design intentions and categories of 3D assistive augmentations.

4.1 AccessMeta: A Metadata of Assistive 3D-Printed Augmentations

We define *"assistive augmentations"* herein as attachments to legacy physical environments, addressing inexplicit barriers in varying contexts. Ever since numerous 3D printing practitioners have open-sourced their creations online (c.f., [10]), many were posted with voluntary textual descriptions with "assistive" to indicate the design intention. Some not originally intended to be assistive missing relative tags could also be used for access but makes shopping through millions of designs by searching exhausting. Navigating options is even more laborious due to ambiguity in language [36]. The structured rules or metadata to categorize assistive augmentations will broaden access to those designs, enabling users to explore easily.

4.1.1 Coding corpus of assistive augmentations. To tackle this, we surveyed large-scale data about designs on Thingiverse [29], defining rules by observation such as retrieving relevant designs for target objects of interest. As our goal is to assist users in searching 3D augmentations based on target objects in mind as approached similarly in prior works [14] and practice (e.g., ThisAble project [67]), we initiated our search with target objects, e.g., "assistive door lever". While the existing categorization and corpus [10] could be useful, designs classified under them do not necessarily represent augmentations. This also applies to CustomizAR taxonomy [36], which primarily focuses on adaptive designs but assistive designs are only a small set. Consequently, we opted not to directly adopt this taxonomy in our corpus formation process.

We selected the initial search keywords of common indoor objects: *door, drawer, cupboard, closet, outlet, light switch, switch, kitchen, utensil, cutlery, knife, spoon, fork, bottle, jar, bag, key, soap, shampoo, dispenser, nail clipper, can, pen, book, spray, phone, laptop, camera, toothbrush, toothpaste, clock, etc.* We started observing the first 50 entries for their affinity defining the corpus. Then we expanded the search, resulting in ~1,600 entries by two sets of keywords overlap. The first and second authors manually annotated their affinity by the common interaction types (Section 3.1), and the last author validated the results for agreement. With iterations and polishing, we define **AccessMeta**, the three high-level categories and their assistive functions, and common keywords and tags (Table 2).

Category	Functions	Common keywords for 3D assistive augmentations
actuation	operation reach	lever/hand extension, grip, mount, opener holder/gripper, string extension
constraint	limit access	cover, guard, protector, lock
indication	visual tactile	label, identifier, tag

Table 2: AccessMeta corpus to categorize 3D assistive augmentations. We found that the majority fall into three categories depending on their desired functions by augmenting real-world objects, often described by common keywords.

(1) Actuation: Reduce motor requirements refers to designs assisting people with operational difficulties (e.g., fine motor impairments, occupied hands) by extending or magnifying parts; including designs that reduce the required strength or alter the needed motion types. Two functions are afforded if augmented: (help) *operation* and *reach*. *Actuation-operation* designs enable alternative operations using other body parts (e.g., elbow-push instead of hands-grab & rotate) or motions or reduce needed power. As an example, a door-knob extension (as in Figure 2b) replaces the grasping-to-rotate with pushing-down. Figure 2d allows other body parts, arm or wrist in this example, for operation, instead of hands that might
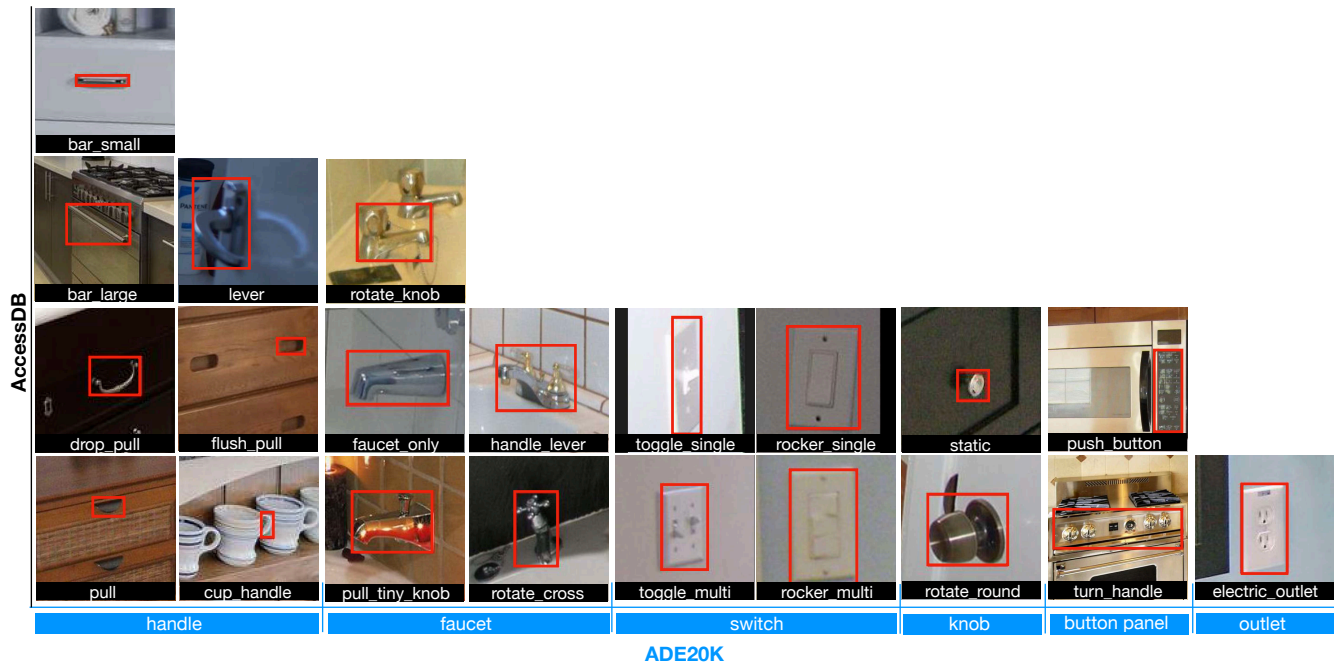
Figure 8: We derive our AccessDB dataset by annotating indoor images from ADE20K dataset [77] with 21 inaccessibility classes. We focus on 6 types of objects (blue-labeled names) which frequently appear to be inaccessible in daily life.

be unavailable at the moment. Another example is a plastic bottle opener [44], which induces the leverage. Different types of pen grips (e.g., [45]) are popular for artists as they degrade necessary wrist-power. *Actuation-reach* designs magnify parts to reach the target. For example, light switch extension [21] is useful for children, people with short stature, or situations where large furniture placed underneath makes access difficult for people using walkers.

(2) Constraint: Prevent operations refers to designs often revert functions of *actuation* designs, preventing operating objects in special contexts (e.g., cabinet lock) for people with cognitive impairments or in child-access/child-proof products. Limiting access is another popular objective in augmentations (e.g., drawer lock [15]) favored by parents, pet owners, and caretakers of the cognitive retreat, especially for safety. Even for those who do not have such impairments, people label identical objects such as a series of wall switches to reduce confusion and misuse. Common target objects contain doors, drawers, wall switches (e.g., lights and garbage disposal), or outlets that are with known risks.

(3) Indication: Furnishing with visual/tactile cues Designs that furnish multi-modal feedback for easy identification of intention, function, or purpose by providing labels (e.g., switch labels, toggling sound); greatly benefit people with sensory impairments. 3D printed tactile graphics have gained acceptance by many people with visual impairments [10]. Built upon those principles, tactile cues provide multi-modal information to help identify functionalities in identical-looking objects, for example, 3D-printed labels in the multi-switch panel [70]. Note that AccessMeta categories are not always mutually exclusive, as one can simultaneously furnish tactile cues and reduce motor requirements.

4.1.2 Assistive 3D Augmentation Dictionary. As a result of design exploration to define AccessMeta, we created an initial dictionary that contains 280 3D-printed augmentations for 52 everyday objects (e.g., handle, door, knob, book, nail clipper, knife, hair dryer, microwave, stove, table, etc.) with potential inaccessibility context, fully annotated with AccessMeta categories. Among 52 common object classes in AccessMeta, we found that 6 classes (i.e., handle, faucet, switch, knob, button panel, and outlet) are significant and difficult to be addressed by existing datasets with indoor scenes (e.g., ADE20K [77], COCO [38]) mainly due to (1) challenges caused by their small size in photos and (2) diverse types of the objects that might pose various kinds of barriers (e.g., door lever vs. knob). Focusing on these 6 classes (which are further divided into 21 inaccessible classes), we construct a new dataset, AccessDB/AccessReal. This dictionary is publicly available at https://access-lens.web.app/.

4.2 AccessDB & AccessReal: Dataset for Inaccessibility Detection

Auto-detecting objects with their semantics and context from camera views (e.g., [50, 55, 68]) can assist visual perception for various interested groups and information processing, e.g., robotic affordance and different types of disability. Automation through a comprehensive dataset that provides a granularity of object classes is critical to infer necessary information from semantics. Yet, predicting contexts from images is more complex than detecting objects and instances; *object attributes* such as shapes (e.g., round, lever, cross-shaped) must relate their *functional properties* (e.g., grip, twist, pinch), to be able to derive their conceptual interaction types. Once interaction types are inferred regarding their visual and functional

AccessDB: training data, ADE20K indoor scenes   **AccessReal**: testing data, modern indoor scenes
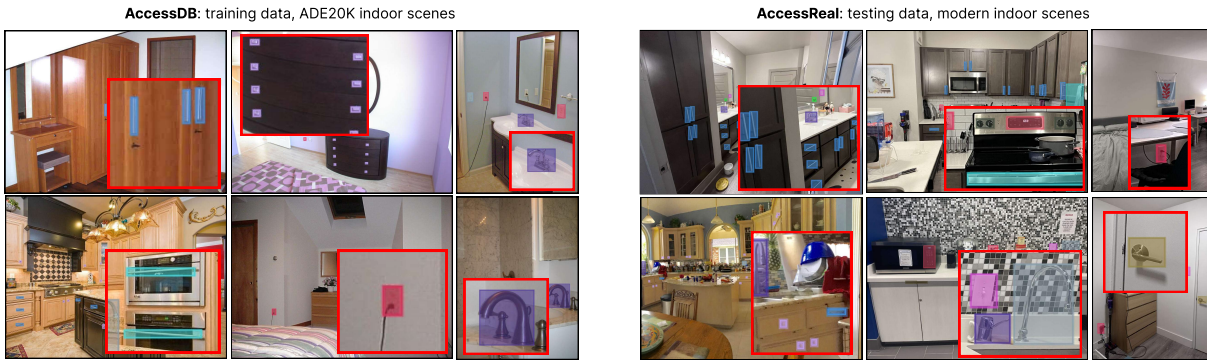
Figure 9: We use the AccessDB (left) and AccessReal (right) datasets to train and evaluate inaccessible-object detectors. Images of AccessDB are sampled from the well-established ADE20K dataset [77] with our re-annotation (cf. Table 3). AccessReal has high-resolution images captured by ourselves from diverse indoor scenes; we annotate these images using the same set of inaccessibility classes. Red boxes are zoom-in regions that contain inaccessible objects.

characteristics, those types can serve as clues to infer the original design intent as well as hidden barriers in various possible contexts. To train and evaluate our developed inaccessibility detector, we construct two datasets: AccessDB and AccessReal. Being built for semantic understanding of objects and their parts, ADE20K offers hierarchical annotations on object classes, such as *closet - door - handle* and *oven - door - handle*. AccessDB presents Inaccessibility Class (IC) to provide a nuanced understanding of diverse barriers that may manifest across various contexts, extracted from six distinct categories in ADE20K: button panels, electrical outlets, faucets, handles, knobs, and switches. The granularity of IC permits the identification of specific accessibility challenges, thus enabling tailored design solutions. Refer to Table 3 in the Appendix A.

Figure 10: An example image (a) from AccessDB with two inaccessible objects: a flat button panel in a stove (b), and a handle into a drawer (c). These objects are often very small in the image, making annotation and detection difficult.

(1) AccessDB is used to train inaccessibility detectors. We derive AccessDB from ADE20K [77], which contains >20k images including diverse indoor scene photos with pixel-level annotations on objects and their parts. We re-annotated objects in ADE20K for 21 predefined inaccessibility classes (ICs) in addition to "unidentifiable" class for extremely small sized parts. We first select scene images sampled from "home", "hotel", "shopping and dining rooms",

and "workplace", but excluded low-resolution images. We focus on 6 object categories that are often inaccessible (Figure 8): handle, faucet, switch, knob, button panel, and electric outlet. Three annotators are HCI experts in assistive designs, and annotators also cross-verify each other's annotations for annotation quality. We obtained 4,976 high-resolution images exhaustively annotated with ICs as illustrated in Figure 10, which appears in extremely small regions of the image, posing a visibility challenge to detectors.

(2) AccessReal. Since AccessDB's images are from the ADE20K dataset which was published five years ago (as of 2023 when this research is conducted), we are motivated to curate a new dataset for evaluation by collecting photos taken in 'modern' indoor scenes. To this end, we take 42 high-resolution photos (mostly 4032×3024) in diverse indoor scenes: bathroom, bedroom, kitchen, living room, and office (cf. Figure 9). We annotate them w.r.t the predefined 21 ICs (see data statistics in Appendix A Table 3), and end up with 428 annotated objects with ICs.

5 EVALUATION

5.1 An End-to-end Pipeline

5.1.1 Participants & Procedure. We conducted a holistic end-to-end study to assess: (1) capturing photos, (2) uploading photos for AI inspections, (3) viewing suggestions to address identified barriers, and (4) physically installing 3D printed results. We recruited six participants (U1-6) from our institution (female=4, male=2, ages 19-30) who have none to limited exposure to accessibility, except for U6 with moderate experience in technology for sign language speakers. Five (U1-5) had little or no prior experience in 3D printing, while U6 had 5+ years of experience in fabrication. None overlapped with the preliminary evaluation study participants. All studies were conducted individually. Participants first freely explored AccessLens, either on mobile or the web. They were asked to upload photos of personal space, and then select as many desired augmentations. Due to time constraints, we printed the chosen augmentations, except for U6 who self-printed. All attached augmentations within their environments by themselves. Participants were asked to take photos and share the installation process, results, and thoughts. We concluded each study session with exit interviews. We took an approach similar to a contextual inquiry, with in-depth observation

Figure 11: Examples of indoor scene photos submitted by case study participants through AccessLens. All participants took photos to show a full coverage of rooms, capturing the details as much as possible. Indoor scenes include: (a-b) bathroom, (c) bedroom, (d) living room, and (e-g) kitchen. (a-g) show bounding boxes overlaid, detected by AccessLens. Participants reported minor detection errors: undetected hair dryer (h) and air fryer misclassified as a toaster (i).

and interviews to gain a robust understanding of user behaviors and their motivation about specific courses of action taken, minimally intervening in the use case. All conversations and responses were transcribed and documented for analysis through coding.

5.1.2 Results & Finding. Participants submitted an average of 3.7 photos/participant, totaling 22 of bathroom, bedroom, living room, and kitchen (e.g., Figure 11).

#1. Easy Photo-taking and Uploading. Although AccessLens did not provide step-by-step instructions and the facilitator minimized intervention, all naturally submitted photos of panoramic views, capturing entire rooms. U5-6 iteratively adapted their photo-shooting strategy, *"From the first try, I saw that the app detected door handles, so I ensured their visibility in subsequent photos"* (U5). None had issues in processing photos and stated it is straightforward.

#2. Learning Accessibility from Adaptation. Before using AccessLens, all participants expressed their lack of confidence in recognizing inaccessibility. U5 guessed that it is possible only when obvious, e.g., seeing someone struggling in person. U1-3 stated they *"had not encountered accessibility challenges themsselves"*, and U4 found it hard *"to view things from the perspective of those with accessibility issues [because I am not disabled]".*

After AccessLens use, we observed elevated confidence and awareness. *"By seeing all the examples and possible solutions in my room, I now have a better understanding of potential issues and how others interact with objects differently from I do"* (U1). U2 found the microwave button pusher [20] eye-opening, since they never imagined that anyone could struggle with such simple pressing. Most participants (U1-4, U6) testified an expansion of their perspectives; *"I never thought outlets or stove buttons [could be inaccessible], since I was expecting more about people who are visually impaired or with [more serious disabilities]. I gained a new perspective that disability is such a large spectrum"* (U3). U4 also stated, *"At first I thought that the challenges would only apply to people with [diagnosed disability, but it applies to] the general population with a variety of issues,*

including injuries, child locks, and having busy hands.", confirming that users learn *"potential contexts"* (U1-2, U6) through recommendations. U5 found being hands-free useful since the steel surfaces tend to become dirty. AccessLens also helped U3 & U6 redefine their experiences; *"I once had a cut on my thumb, which made squeezing the toothpaste very difficult. Toothpaste squeezer seems useful (in such situations) but also on a daily basis too"* (U6).

#3. Perceived Accuracy of Detection. All participants found the automated detection accurate, expressing confidence in interpreting the results. U3 was concerned about messy rooms but was impressed by the detector performance that captured objects successfully even from cluttered scenes. U6 found that even a small reflection of the door knob in a mirror was correctly detected. AccessLens was thought accurate only except for U1's hair dryer, possibly due to its uncommon design (Figure 11h), and U4's air fryer is seen as a toaster (Figure 11i). All were thought minor and did not affect participants' trust in overall detection results.

#4. AccessMeta and Dictionary Supporting Exploration. Participants appreciated AccessLens' presentations, organized by the detected objects and related issues with AccessMeta. Participants (U2-3, U5) found the dictionary explorer, which shows all possible designs useful. *"Before reading the dictionary, I was not aware of child safety and how they related to accessibility, but the dictionary helped me learn potentially dangerous aspects of objects and how to mitigate them"* (U1). U3 perceived the variety of the dictionary as very useful for browsing especially *"when moving to a new place, remodeling, or choosing new appliances".* U6 imagined augmenting standard spaces with various needs; *"The standard apartment's equipment is not designed for specific needs. People will find it very useful to augment their everyday environment with specific needs in mind".*

#5. Different Motivations to Adopt AccessMeta Recommendations. Each participant selected 2-4 augmentations, such as a hands-free opener for large door handles, electric outlet covers, jar openers, stove knob protectors, etc (example retrofitting results
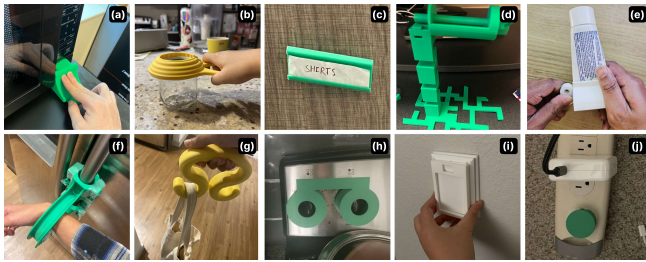
Figure 12: Retrofitting 3D-printed augmentations by study participants: (a) microwave opener, (b) jar opener, (c) drawer label holder, (d-e) toothpaste squeezer, (f) hands-free door opener, (g) bag holder, (h) stove knob cover, (i-j) outlet cover.

seen in Figure 12). Their selection criteria varied: frequency of use (U1, U6), assistance when alone (U2), safety (U3, U5), practicality, and sheer interest (U4). Some could still find useful designs through an inductive process, not necessarily having the images; *"I know my parents or grandparents struggle using, such as a toenail clipper as they don't have enough back flexibility. It's nice to have the option to look at suggestions [without having the images] of their houses"* (U2). We imagine AccessLens' advanced feature for expanded recommendations. If the contextual disabilities are known through the user's previous choices of recommended adaptations, AccessLens can fetch common objects that present similar barriers.

#6. Low-cost Upgrades through Retrofitting but Need to Handle Uncertainty. Participants found 3D-printed upgrades easy and cost-effective. All were able to install the augmentations without any help and did not face major difficulties, spending a maximum of a few minutes when designs required assembly. Many designs on Thingiverse are versatile and modular, often in standard dimensions or using screws for a tight fit. Participants found standalone designs (e.g., bag holders, knob covers) were easy to utilize. For example, U3 found that the stove knob lock fit perfectly, and found it useful for safety when children or cats are around. With assembly, participants were actively involved in the adaptation. U1 found that the microwave door opener [20] is slightly taller, so they tilted the microwave up to match the height. U4 and U5 did not have screws to put parts of the hands-free door opener [72], but still made it work by installing it using tape. For designs that need assembly, three participants (U1-3) thought having a step-by-step guide would be beneficial. While all successfully adopted designs, some reported dimensional challenges; U3's outlet covers did not fit so they had to put it over without fixation. U5's hands-free fridge opener was loose and slid, failing to stay at arm height. We consider integrating well-established customization tools focusing on a fit, e.g., [24, 31] and auto-measurement [36].

#6. Additional Suggestions. Overall, participants were satisfied and willing to continue using AccessLens. Three participants (U3-5) suggested a detailed description for augmentations clarifying the functionality and objectives on the app without redirecting to the design page. U1-2 and U4 also mentioned that showing the required materials (e.g., screws, tape, clips) would be helpful for users to make choices based on complexity and material availability. U6 also hoped to see an animated preview of how the augmentation could change the interaction.

5.2 Expert Feedback about User Experience

5.2.1 Participants. The expert feedback session was conducted to understand how AccessLens can support users to raise awareness about accessibility. We engaged two professionals (E1-2) with 10+ years of expertise in accessibility research and teaching access computing. E1's expertise lies in robotics for people with movement disabilities and/or chronic conditions (e.g., people with Parkinson's disease, and freezing of gait), and E2's expertise is in assistive visual perception for the visually impaired through systems for human-AI interaction. We sought their qualitative opinions about various topics of interest: user engagement, system functionality, empowerment in decision-making, alignment with standards, usability, potential impact, and future developments.

5.2.2 Findings. Both acknowledged the tool's diverse and relevant suggestions, particularly for *"raising awareness of accessibility issues, aiding those without specialized accessibility knowledge"* (E1). Yet, E1 expressed concerns about non-experts due to the absence of a clear description of relevant accessibility issues for diagnosed disabilities. While the system provides real examples and suggestions for environmental modification facilitating users' perception of various possible contexts indirectly, it lacks *"explicit explanations"*, potentially hindering informed decision-making. E2 commented about possible design conflicts, *"if multiple people residing in the space with different accessibility needs, solutions could be in conflict with each other, or the design needs to be combined to satisfy multiple needs"*. AccessLens needs more targeted customization and alignment with *public accessibility standards*, E1 added. Similarly, while appreciating the system's ability to identify numerous relevant objects, E2 suggested incorporating design parameters, including configuration/layout of the environment (e.g., the width of a hallway) and interaction/spacing between objects (e.g., the distance between switch and floor), which we find incorporating physical assertion of adaptive designs [24] critical. E1 sees long-term benefits, particularly for growing 3D printing communities but with limited accessibility knowledge. E2 also proposed allowing users to input disability types to prioritize suggestions and emphasize the importance of customizing solutions in mind. E2 imagined crowdsourcing for more examples and an onboarding feature for new users to enhance utility. In summary, both experts recognize AccessLens's potential to engage inexperienced users. Encompassing customization support to accommodate various physical dynamics, guidance, and user-defined disability prioritization at the input stage can further improve AccessLens.

5.3 AccessMeta's Acceptance

5.3.1 Procedure. To assess the acceptance of AccessMeta, we conducted an independent study on human annotators' perception and consensus on AccessMeta and a fully annotated dictionary of 280 3D augmentations by the research team. Using Amazon Mechanical Turk, we designed tasks to assess how well the general public understands AccessMeta's classification criteria. In each HIT, annotators engage with one 3D augmentation and categorize it under one of the three high-level categories from AccessMeta: 'actuation', 'constraint', and 'indication'. These categories are further organized into five sub-categories: 'actuation-reach', 'actuation-operation', 'constraint', 'indication-visual', and 'indication-tactile'.
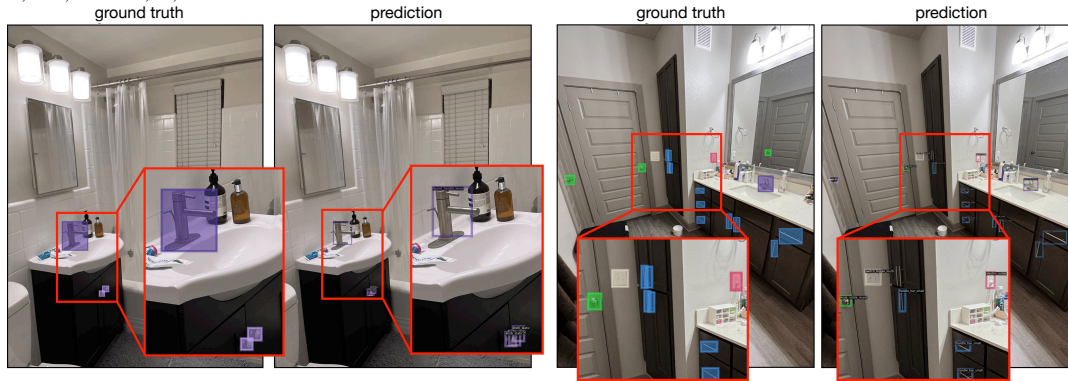
Figure 13: Ground-truth and detection results of our inaccessible-object detector on two example images in AccessReal. For brevity, we omit IC labels (and detection confidence scores) in ground truth but present only labels for detection boxes. A visual examination of the results reveals that our detector exhibits a decent capability for identifying inaccessible objects.

An additional 'others' option was provided for a custom label, if any. Annotators opened a provided URL (e.g., Thingiverse) and chose the label(s) that best describe the augmentation. To avoid potential bias, we did not provide any image references. Instead, annotators are provided textual descriptions from AccessMeta. We consider a HIT submission acceptable in acceptable scenarios: an annotator (1) correctly identifies the specific label (e.g., 'actuation-reach'), (2) chooses a subcategory under the correct high-level category (e.g., selecting 'actuation-operation' for an 'actuation' design), (3) chooses multiple labels within the correct high-level category, and (4) selects 'others' with a reasonable custom label.

Submitted HITs were first reviewed by the second author and were subject to rejection only when they fell under four cases: (1) *all* annotations provided by a single annotator for different design entries were identical and incorrect (all HITs would be rejected); (2) an annotator selected 'others' but provided irrelevant tags such as too generic comments ('good design'), unrelated phrases ('We and our 814 partners'), or simply copied the full title or description of the design; (3) all responses submitted by a single annotator were incorrect and completed in less than 40 seconds (threshold decided from the test run), which indicates insufficient time to complete the task; (4) a single annotator submitted more than 100 HITs, any responses beyond the 100-HIT limit would be rejected to ensure diversity in results. Results were again verified by the first author. N=515 HITs were rejected and republished for re-annotation. A worker was paid $0.05 per HIT, and one worker submitted 16.8 annotations on average.

5.3.2 Results & Findings. Three different annotations were collected for each of the 280 designs, eventually obtaining 839 valid annotations from 83 workers. The median completion time was 6.8 minutes (8 sec. to 30 min., std = 6.8 min.)

Acceptability. If workers' annotations matched the ground truths of three main classes, they were marked as success, otherwise, failure. Accuracy was analyzed by the ratio of correct annotations over total annotations obtained (N=839) across 280 designs. Annotators showed 83% match (N=697), implying fair acceptance of AccessMeta. For about 20% of correct annotations, workers' selection of subcategories varied within a category, e.g., 'actuation-reach' instead of 'actuation-operation', possibly due to the versatile nature of assistive designs. As discussed earlier, AccessMeta subcategories are not always mutually exclusive. For instance, tactile indications

often provide visual cues, and extensions to help reach items can also facilitate alternative or smoother operation.

Category Expansion by Annotator-Adaptation. About 98% of annotations were made from AccessMeta categories. Despite not many (1.8%), 10 workers selected the 'others' option for 13 designs. Three new classes emerged, mostly for designs labeled as 'actuation-operation' (e.g., hands-free book holder [13], ziploc back holder [73], cup holder attachable to the sofa [69]): 'holder' (N=6), 'stabilizer' (N=2), and 'support' (N=2). Annotators also suggested 'protector' (N=2) and 'safety' (N=1) for child-proof designs—a child finger protector for drawers [12] and a sharp corner protector [52], respectively, which are currently defined as 'constraint'. Growing in complexity with diverse contexts and objects, we perceive AccessMeta to serve as a platform to expand through the collective input for more diverse & inclusive classifications. Future work could involve mechanisms for reports/suggestions from stakeholders and designers for adaptive solutions.

5.4 AccessDB Detector Performance

Our approach allows adapting any state-of-the-art detector architectures (e.g., GroundingDINO [40] and RetinaNet [37]; details in Appendix C). Figure 13 displays example detection results on AccessReal images, showing good qualitative performance in detecting small inaccessible objects. In evaluation, (Section 5.1), all participants showed solid trust in our detector's performance. Detection result visualizations for sample images in AccessReal (Figure 13) also show that the detector accurately captures small objects. AccessDB and AccessReal datasets are open-sourced at https://access-lens.web.app/ to foster future research. While our work used one of the state-of-arts, any modules can be trained on our dataset. For technical specifications, refer to Appendix C.

6 DISCUSSION & FUTURE WORK

6.1 Collective Disability Accommodations

Engaging with a building ADA coordinator at our institution sheds light on a collective effort in identifying/reporting. The ADA coordinator admitted that many staff lack accessibility expertise, so they hire external accessibility specialists to address issues on demand. Encouraging citizen science within our initiative could mirror successful collective intelligence models like Project Sidewalk [64]. By

adopting a reporting system where individuals contribute to accessibility assessment within commons, accommodating potentially inaccessible physical environments but have *not yet* discovered by people with diagnosed disabilities before they encounter barriers. Experts' recommendations to input disability types and validate possible conflicts can be applied to seek AccessLens at scale.

6.2 Is AccessLens A Disability Dongle?

The term 'disability dongles' has emerged to criticize endeavors that employ innovative technologies but fail to address genuine accessibility needs [41], often targeting industry products that exploit accessibility concepts for superficial gains. AccessLens builds upon prior work on understanding and addressing existing issues of recognizing accessibility barriers [68]. Our approach utilizes the state-of-the-art technologies that are becoming more and more available to innovate an individual's life (e.g., [6]), thereby assisting users make informed decision-making. Moving one step forward, AccessLens broadens its impact to a wider audience who do not experience diagnosed disability, provoking discourse about disabled contexts [48]. 3D printed and/or DIY solutions already become the major efforts made by numerous disabled selves, stakeholders, and altruistic enthusiasts [10, 14]. As AccessLens promises to incorporate more options for store-bought solutions and the industrial design industry (e.g., [55]), we anticipate more collaborative efforts across disciplines to advance people's quality of life using technology as we detail in the following section.

6.3 Expanding to 3rd-Party Solutions

AccessMeta links the object types with their needed interaction, seeking solutions that might alter interaction types (e.g., grab-rotate-to-open vs. push-open). Once detected, we see the future of AccessMeta and the dictionary expanding the search for similarly-functioning 3rd-party alternatives, such as buying door lever replacements from hardware stores or online markets. While some simple replacements like doorknobs might be as cheap as 3D printing, more complex fixtures such as refrigerator handles (as in Figure 12) are not trivial, necessitating the disassembly or replacement of the whole appliances. Although our study participants agreed on the less mental burdens with AccessLens recommendations, some were inclined towards store-bought products as they have gone through market testing (U3), given their perceived affordability and time cost for customization (U5). As AccessLens provides direct recommendations compared to *"for store-bought ones, I might have to look for products on my own that solve the highlighted challenge for detected objects"* (U2), offering users more options upon various rationale, control for materials (U5), easy-fix and remix (U6).

6.4 3D Model Customization

3D printing is a promising solution for custom adaptive interfaces to meet unique needs. One notable example is auto-filling numerics into parametric 3D designs [36] and in creating various branches of augmentations upon user's needs to adapt common household items [14]. The current AccessLens prioritizes inaccessibility detection and assistive augmentation recommendations. As our work has been focusing on increasing awareness and low-cost solutions, dealing with fit [31] and parametric customization was considered

orthogonal. However, we recognize the potential synergy with existing works facilitating customization (e.g., [24]), starting from the auto-detection and selection of a suitable design and culminating in real-world applications. We can further empower individuals to take proactive steps toward creating inclusive environments.

6.5 Expanding AccessDB & AccessReal Dataset, Populating AccessMeta

This work provides two challenging datasets, AcessDB and AccessReal for inaccessibility detectors. Communities' interest in inclusive designs has grown, and advances to automate everyday surroundings (e.g., smart switches, thermostats with touch screens) create new challenges; touch screens often lack tactile feedback for people with visual impairments and can present more challenges for the elderly). To scale the dataset, this work elaborated on the re-annotation strategy of AccessDB in detail at our dataset website. We believe that the AccessMeta pipeline should remain open-ended and adaptable to accommodate emerging needs and novel designs. One approach to expanding the AccessMeta pipeline is involving a community in reporting problems and suggesting additional metadata categories. We can ensure that the system remains responsive to real-world needs, identifying new challenges.

6.6 Can AccessLens Promote Altruism?

We envision the use of AccessLens will help people become more aware of implicit inaccessibility and more actively engaged in improving access in public spaces, such as lecture rooms and shared dormitory community rooms. We have not observed positive behavioral changes in participants beyond the lab. We plan to conduct a deployment study to evaluate whether AccessLens raises people's awareness and encourages collective actions, similar to how altruism motivates voluntary sharing of designers online for free. Expert interviews from diverse domains, including HCI, accessibility, visualization, and citizen science, will be conducted to critique the user interface and study design systematically, ensuring unbiased evaluation of AccessLens compared to other existing tools.

7 CONCLUSION

AccessLens provides an end-user tool that helps users without diagnosed disabilities or prior experiences in accessibility assess the accessibility challenges. We adopted object detection techniques to train inaccessible-object detectors on our novel dataset AccessDB. On our collected dataset AccessReal which consists of images of modern indoor scenes, we show that our detector can detect inaccessible-objects well. We designed AccessMeta to link inaccessibility classes to keywords of 3D assistive augmentations. Through two rounds of holistic evaluation with inexperienced users, we demonstrate the effectiveness of AccessLens in raising awareness and proactiveness in improving indoor accessibility.

REFERENCES

[1] 2020. What is Universal Design? https://universaldesign.ie/what-is-universal-design/. Accessed: 11/14/2023.

[2] Phil Ammirato, Cheng-Yang Fu, Mykhailo Shvets, Jana Kosecka, and Alexander C Berg. 2018. Target driven instance detection. *arXiv preprint arXiv:1803.04610* (2018).

[3] Abul Al Arabi, Jiahao Li, Xiang'Anthony Chen, and Jeeeun Kim. 2022. Mobiot: Augmenting Everyday Objects into Moving IoT Devices Using 3D Printed Attachments Generated by Demonstration. In *CHI Conference on Human Factors in Computing Systems*. 1–14.

[4] Adrienne Asch and Henry McCarthy. 2003. Infusing disability issues into the psychology curriculum. (2003).

[5] Daniel Ashbrook, Shitao Stan Guo, and Alan Lambie. 2016. Towards augmented fabrication: Combining fabricated and existing objects. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*. 1510–1518.

[6] Patrick Baudisch, Stefanie Mueller, et al. 2017. Personal fabrication. *Foundations and Trends® in Human–Computer Interaction* 10, 3–4 (2017), 165–293.

[7] Cynthia L Bennett, Erin Brady, and Stacy M Branham. 2018. Interdependence as a frame for assistive technology research and design. In *Proceedings of the 20th international acm sigaccess conference on computers and accessibility*. 161–173.

[8] Cynthia L Bennett and Daniela K Rosner. 2019. The promise of empathy: Design, disability, and knowing the" other". In *Proceedings of the 2019 CHI conference on human factors in computing systems*. 1–13.

[9] Alexander Berman, Francis Quek, Robert Woodward, Osazuwa Okundaye, and Jeeeun Kim. 2020. "Anyone Can Print": Supporting Collaborations with 3D Printing Services to Empower Broader Participation in Personal Fabrication. In *Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society*. 1–13.

[10] Erin Buehler, Stacy Branham, Abdullah Ali, Jeremy J Chang, Megan Kelly Hofmann, Amy Hurst, and Shaun K Kane. 2015. Sharing is caring: Assistive technology designs on thingiverse. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. 525–534.

[11] Marco Buzzelli, Alessio Albé, and Gianluigi Ciocca. 2020. A Vision-Based System for Monitoring Elderly People at Home. *Applied Sciences* 10, 1 (2020). https://doi.org/10.3390/app10010374

[12] celiktse (Thingiverse User). December, 2018. child finger protector. https://www.thingiverse.com/thing:3286025. Accessed: 12/9/2023.

[13] cfisch06 (Thingiverse User). March, 2020. Hands-Free Book Holder. https://www.thingiverse.com/thing:4232488. Accessed: 12/9/2023.

[14] Xiang'Anthony' Chen, Jeeeun Kim, Jennifer Mankoff, Tovi Grossman, Stelian Coros, and Scott E Hudson. 2016. Reprise: A design tool for specifying, generating, and customizing 3D printable adaptations on everyday objects. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*. 29–39.

[15] davem80 (Thingiverse user). 2018. Magnetic Drawer Lock (child safety). https://www.thingiverse.com/thing:3075157. Accessed: 4/3/2023.

[16] U.S. department of Justice Civil Rights Division. 2010. 2010 ADA Standards for Accessible Design. https://www.ada.gov/law-and-regs/design-standards/2010-stds/. Accessed: 3/23/2023.

[17] Debidatta Dwibedi, Ishan Misra, and Martial Hebert. 2017. Cut, paste and learn: Surprisingly easy synthesis for instance detection. In *Proceedings of the IEEE international conference on computer vision*. 1301–1310.

[18] Philip Easom, Ahmed Bouridane, Feiyu Qiang, Li Zhang, Carolyn Downs, and Richard M. Jiang. 2020. In-House Deep Environmental Sentience for Smart Homecare Solutions Toward Ageing Society. *2020 International Conference on Machine Learning and Cybernetics (ICMLC)* (2020), 261–266.

[19] Yasmine El-Glaly, Weishi Shi, Samuel Malachowsky, Qi Yu, and Daniel E Krutz. 2020. Presenting and evaluating the impact of experiential learning in computing accessibility education. In *Proceedings of the ACM/IEEE 42nd International Conference on Software Engineering: Software Engineering Education and Training*. 49–60.

[20] ezra_reynolds (Thingiverse User). January, 2015. Microwave Door Opener. https://www.thingiverse.com/thing:4232342. Accessed: 12/9/2023.

[21] fastryan (Thingiverse user). 2020. Toddler Light Switch Extension. https://www.thingiverse.com/thing:4098133. Accessed: 4/3/2023.

[22] Anhong Guo, Jeeeun Kim, Xiang 'Anthony' Chen, Tom Yeh, Scott E. Hudson, Jennifer Mankoff, and Jeffrey P. Bigham. 2017. Facade: Auto-generating Tactile Interfaces to Appliances. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '17)*. ACM, New York, NY, USA, 5826–5838. https://doi.org/10.1145/3025453.3025845

[23] Tomáš Hodaň, Vibhav Vineet, Ran Gal, Emanuel Shalev, Jon Hanzelka, Treb Connell, Pedro Urbina, Sudipta N Sinha, and Brian Guenter. 2019. Photorealistic image synthesis for object instance detection. In *IEEE international conference on image processing (ICIP)*.

[24] Megan Hofmann, Gabriella Hann, Scott E Hudson, and Jennifer Mankoff. 2018. Greater than the sum of its PARTs: Expressing and reusing design intent in 3D models. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–12.

[25] Megan Hofmann, Devva Kasnitz, Jennifer Mankoff, and Cynthia L Bennett. 2020. Living disability theory: Reflections on access, research, and design. In *Proceedings of the 22nd International ACM SIGACCESS Conference on Computers and Accessibility*. 1–13.

[26] Derek Hoiem, Santosh K Divvala, and James H Hays. 2009. Pascal VOC 2008 challenge. *World Literature Today* 24 (2009).

[27] Sarah Horton. 2021. Empathy cannot sustain action in technology accessibility. *Frontiers in Computer Science* 3 (2021), 617044.

[28] IBC. 2021. International Building Code (IBC) Chapter 11 Accessibility. https://codes.iccsafe.org/content/IBC2018/chapter-11-accessibility. Accessed: 4/3/2023.

[29] MakerBot Industries. 2008. Thingiverse. https://www.thingiverse.com/. Accessed: 9/10/2022.

[30] ippe (Thingiverse User). 2020. Hands-free door handle (Coronavirus prevention). https://www.thingiverse.com/thing:4225872. Accessed: 3/22/2023.

[31] Jeeeun Kim, Anhong Guo, Tom Yeh, Scott E Hudson, and Jennifer Mankoff. 2017. Understanding uncertainty in measurement and accommodating its impact in 3D modeling and printing. In *Proceedings of the 2017 conference on designing interactive systems*. 1067–1078.

[32] Masaki Kuribayashi, Tatsuya Ishihara, Daisuke Sato, Jayakorn Vongkulbhisal, Karnik Ram, Seita Kayukawa, Hironobu Takagi, Shigeo Morishima, and Chieko Asakawa. 2023. PathFinder: Designing a Map-less Navigation System for Blind People in Unfamiliar Buildings. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–16.

[33] Waag Future Lab. 2019. MakeHealth. https://waag.org/en/project/makehealth/. (Accessed on 09/08/2023).

[34] Chengshu Li, Ruohan Zhang, Josiah Wong, Cem Gokmen, Sanjana Srivastava, Roberto Martín-Martín, Chen Wang, Gabrael Levine, Michael Lingelbach, Jiankai Sun, et al. 2023. Behavior-1k: A benchmark for embodied ai with 1,000 everyday activities and realistic simulation. In *Conference on Robot Learning*. PMLR, 80–93.

[35] Jiahao Li, Jeeeun Kim, and Xiang'Anthony' Chen. 2019. Robiot: A design tool for actuating everyday objects with automatically generated 3D printable mechanisms. In *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology*. 673–685.

[36] Chen Liang, Anhong Guo, and Jeeeun Kim. 2022. CustomizAR: Facilitating Interactive Exploration and Measurement of Adaptive 3D Designs. In *Designing Interactive Systems Conference* (Virtual Event, Australia) *(DIS '22)*. Association for Computing Machinery, New York, NY, USA, 898–912. https://doi.org/10.1145/3532106.3533561

[37] Tsung-Yi Lin, Priya Goyal, Ross Girshick, Kaiming He, and Piotr Dollár. 2017. Focal loss for dense object detection. In *Proceedings of the IEEE international conference on computer vision*. 2980–2988.

[38] Tsung-Yi Lin, Michael Maire, Serge Belongie, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollár, and C Lawrence Zitnick. 2014. Microsoft coco: Common objects in context. In *Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part V 13*. Springer, 740–755.

[39] Tsung-Yi Lin, Michael Maire, Serge J. Belongie, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollár, and C. Lawrence Zitnick. 2014. Microsoft COCO: Common Objects in Context. In *European Conference on Computer Vision*.

[40] Shilong Liu, Zhaoyang Zeng, Tianhe Ren, Feng Li, Hao Zhang, Jie Yang, Chunyuan Li, Jianwei Yang, Hang Su, Jun Zhu, et al. 2023. Grounding dino: Marrying dino with grounded pre-training for open-set object detection. *arXiv preprint arXiv:2303.05499* (2023).

[41] Rua Williams Liz Jackson, Alex Haagaard. April, 2022. Disability Dongle. https://blog.castac.org/2022/04/disability-dongle/. Accessed: 2/8/2024.

[42] Zelun Luo, Jun-Ting Hsieh, Niranjan Balachandar, Serena Yeung, Guido Pusiol, Jay S. Luxenberg, Grace Li, Li-Jia Li, Arnold Milstein, and Li Fei-Fei. 2018. Vision-Based Descriptive Analytics of Seniors ' Daily Activities for Long-Term Health Monitoring.

[43] Zelun Luo, Alisha Rege, Guido Pusiol, Arnold Milstein, Li Fei-Fei, and Norman Lance Downing. 2017. Computer Vision-based Approach to Maintain Independent Living for Seniors. In *American Medical Informatics Association Annual Symposium*.

[44] makersmakingchange (Thingiverse user). 2018. Bottle Opener. https://www.thingiverse.com/thing:2801157. Accessed: 4/3/2023.

[45] makersmakingchange (Thingiverse user). 2018. Pen Ball. https://www.thingiverse.com/thing:2810069. Accessed: 4/3/2023.

[46] Mike Mankin. 2020. My Day in a Wheelchair – How Experience Leads to Empathy. https://www.linkedin.com/pulse/my-day-wheelchair-how-experience-leads-empathy-michael-mankin/.

[47] Jean-Philippe Mercier, Mathieu Garon, Philippe Giguere, and Jean-Francois Lalonde. 2021. Deep template-based object instance detection. In *WACV*.

[48] Microsoft. 2018. Microsoft inclusive design. https://www.microsoft.com/design/inclusive/. Accessed: 8/28/2022.

[49] Microsoft. 2023. An intro to Inclusive Design | Microsoft Inclusive Design. https://youtu.be/42RojZSB0Yg?si=R5oyFF32yZVOWCfb.

[50] Alexandros Mitsou, Dimitra-Christina C. Koutsiou, Dimitrios E. Diamantis, Theodoros Psallidas, George Dimas, Michael Vasilakakis, Panagiotis Kalozoumis, Evaggelos Spyrou, Stavros J. Perantonis, Artur Krukowski, and Dimitris K. Iakovidis. 2022. ENORASI Assistive Computer Vision-Based System for the Visually Impaired: A User Evaluation Study. In *Proceedings of the 15th International Conference on PErvasive Technologies Related to Assistive Environments* (Corfu, Greece) *(PETRA '22)*. Association for Computing Machinery, New York, NY, USA, 668–677. https://doi.org/10.1145/3529190.3534784

[51] Kaichun Mo, Shilin Zhu, Angel X Chang, Li Yi, Subarna Tripathi, Leonidas J Guibas, and Hao Su. 2019. Partnet: A large-scale benchmark for fine-grained and hierarchical part-level 3d object understanding. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 909–918.

[52] motel (Thingiverse User). November, 2018. Corner protector for 8mm thick glass table (Child protection). https://www.thingiverse.com/thing:3214474. Accessed: 12/9/2023.

[53] Michelle R Nario-Redmond, Dobromir Gospodinov, and Angela Cobb. 2017. Crip for a day: The unintended negative consequences of disability simulations. *Rehabilitation psychology* 62, 3 (2017), 324.

[54] Ben Niu and Gang Tan. 2014. RockJIT: Securing just-in-time compilation using modular control-flow integrity. In *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security*. 1317–1328.

[55] American Association of Retired Persons (AARP). 2020. HomeFit AR. https://apps.apple.com/us/app/homefit-ar/id1513619492?platform=iphone. Accessed: 4/3/2023.

[56] American Association of Retired Persons (AARP). 2022. AARP Homefit Guide. https://www.aarp.org/livable-communities/housing/info-2020/homefit-guide.html. Accessed: 3/25/2023.

[57] Cynthia Putnam, Maria Dahman, Emma Rose, Jinghui Cheng, and Glenn Bradford. 2015. Teaching accessibility, learning empathy. , 333–334 pages.

[58] Cynthia Putnam, Maria Dahman, Emma Rose, Jinghui Cheng, and Glenn Bradford. 2015. Teaching accessibility, learning empathy. In *Proceedings of the 17th International ACM SIGACCESS Conference on Computers & Accessibility*. 333–334.

[59] Ariadna Quattoni and Antonio Torralba. 2009. Recognizing indoor scenes. In *2009 IEEE conference on computer vision and pattern recognition*. IEEE, 413–420.

[60] Janet Reid. 2009. Blindfolds teach empathy for visually impaired. https://www2.ljworld.com/news/2009/jun/12/blindfolds-teach-empathy-visually-impaired/?city_local.

[61] Stefanie Reid. 2019. Why accessible design is for everyone. https://www.ted.com/talks/stefanie_reid_why_accessible_design_is_for_everyone.

[62] Mike Roberts, Jason Ramapuram, Anurag Ranjan, Atulit Kumar, Miguel Angel Bautista, Nathan Paczan, Russ Webb, and Joshua M Susskind. 2021. Hypersim: A photorealistic synthetic dataset for holistic indoor scene understanding. In *Proceedings of the IEEE/CVF international conference on computer vision*. 10912–10922.

[63] Elise Roy. 2015. Why accessible design is for everyone. https://www.ted.com/talks/elise_roy_when_we_design_for_disability_we_all_benefit?language=en.

[64] Manaswi Saha, Michael Saugstad, Hanuma Teja Maddali, Aileen Zeng, Ryan Holland, Steven Bower, Aditya Dash, Sage Chen, Anthony Li, Kotaro Hara, et al. 2019. Project sidewalk: A web-based crowdsourcing tool for collecting sidewalk accessibility data at scale. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–14.

[65] Neil Savage et al. 2022. Robots rise to meet the challenge of caring for old people. *Nature* 601, 7893 (2022), 8–10.

[66] Shuran Song, Samuel P Lichtenberg, and Jianxiong Xiao. 2015. Sun rgb-d: A rgb-d scene understanding benchmark suite. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 567–576.

[67] University Design Style. September 2018. ThisAbles | Assistive Technology Devices For Ikea Products. https://www.universaldesignstyle.com/thisables-assistive-technology-devices-for-ikea-products/. Accessed: 4/4/2023.

[68] Xia Su, Kaiming Cheng, Han Zhang, Jaewook Lee, and Jon E Froehlich. 2022. Towards Semi-automatic Detection and Localization of Indoor Accessibility Issues using Mobile Depth Scanning and Computer Vision. *arXiv preprint arXiv:2210.02533* (2022).

[69] Cal_L (Thingiverse User). August, 2019. Ikea Ektorp Sofa Cup Holder. https://www.thingiverse.com/thing:3795538. Accessed: 12/9/2023.

[70] FuzzyOrange (Thingiverse user). 2021. Switchplate Identifiers. https://www.thingiverse.com/thing:4924800. Accessed: 4/3/2023.

[71] Mister_G (Thingiverse User). 2015. E Z Open Door Lever Adaptor. https://www.thingiverse.com/thing:1094505. Accessed: 3/22/2023.

[72] PeeKayFr (Thingiverse User). March, 2020. Circular hands-free opener for front door bar. https://www.thingiverse.com/thing:4232342. Accessed: 12/9/2023.

[73] RC64nut (Thingiverse User). December, 2015. Ziploc Bag Holder. https://www.thingiverse.com/thing:1228294. Accessed: 12/9/2023.

[74] Xiyue Wang, Seita Kayukawa, Hironobu Takagi, and Chieko Asakawa. 2022. BentoMuseum: 3D and Layered Interactive Museum Map for Blind Visitors. In *Proceedings of the 24th International ACM SIGACCESS Conference on Computers and Accessibility*. 1–14.

[75] Yuxin Wu, Alexander Kirillov, Francisco Massa, Wan-Yen Lo, and Ross Girshick. 2019. Detectron2. https://github.com/facebookresearch/detectron2.

[76] Fei Xia, Amir R Zamir, Zhiyang He, Alexander Sax, Jitendra Malik, and Silvio Savarese. 2018. Gibson env: Real-world perception for embodied agents. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 9068–9079.

[77] Bolei Zhou, Hang Zhao, Xavier Puig, Sanja Fidler, Adela Barriuso, and Antonio Torralba. 2017. Scene parsing through ade20k dataset. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 633–641.

A DATASET DETAILS

AccessDB/Real comprises around 10k re-annotated objects across 21 ICs. Further details regarding the breakdown of our dataset can be found in Table 3.

id	inaccessibility class	AccessDB	AccessReal
1	button_panel_push_buttons	83	14
2	button_panel_turn_handle	165	8
3	electric_outlet	1,382	33
4	faucet_faucet_only	169	3
5	faucet_handle_lever	351	13
6	faucet_pull_tiny_knob	29	0
7	faucet_rotate_cross	86	0
8	faucet_rotate_knob	96	0
9	handle_bar_large	375	19
10	handle_bar_small	1,712	191
11	handle_cup_handle	243	31
12	handle_drop_pull	491	0
13	handle_flush_pull	43	0
14	handle_lever	211	10
15	handle_pull	289	14
16	knob_rotate_round	205	26
17	knob_static	3,026	38
18	switch_rocker_multi	84	3
19	switch_rocker_single	57	4
20	switch_toggle_multi	103	8
21	switch_toggle_single	115	13
22	unidentifiable	724	0
	total	**10,039**	**428**

Table 3: Counts of annotated objects per inaccessibility classes in AccessDB and AccessReal datasets. There are 21 inaccessibility classes plus an "unidentifiable". AccessDB and AccessReal contain 2,388 and 42 indoor scene images, respectively. We use AccessDB for training and validation, and AccessReal as the testing set for evaluation.

B EXAMPLE ASSISTIVE AUGMENTATIONS

We introduce example augmentations with possible user walk-through of AccessLens in various user contexts in Figure 14.

Scenario #1: Home Adjustment for a Mom. Mark's wife struggles to care for a 6-month-old baby and do housework. Mark wanted to upgrade his home. He scanned the rooms using AccessLens to get recommendations for common objects such as doorknobs (Figure 14 1a), water faucets, and lower drawers (1e-f). It also proposed an arm-activated handle (1a) and a foot-open door handle (1b). It suggested an one-handed hand soap dispenser which he did not notice was hard to use but useful if it applied to the baby lotion bottle for one-hand dispense (1c). Mark now understands what could be inaccessible with arms occupied by a baby.
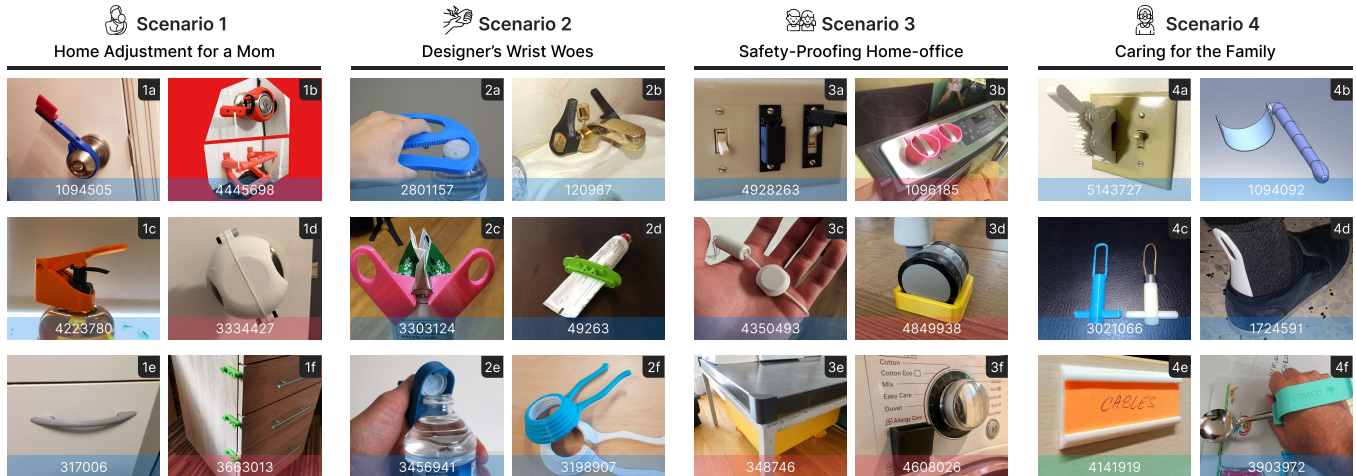
Figure 14: Augmentations recommended by AccessLens for four example scenarios: (1) home adjustment for a new mom, (2) designers who has chronic wrist pain, (3) safe-proofing home office, and (4) caring for the family member who is an older adult. Each design has a *thing_id* at the bottom label, and the design page locates at https://www.thingiverse.com/thing:*thing_id*. Labels indicate blue: actuation, red: constraint, green: indication.

Scenario #2: Designer's Wrist Woes. Enter Kathy, a designer grappling with chronic wrist pain that hindered daily tasks. AccessLens scrutinized her kitchen, finding jar lids (2f), plastic bottles (2a, c), and milk cartons (2c) were known to be inaccessible. Provided with custom-designed openers tailored to each item, she feels relief for her weakened wrists. Kathy also opted for 3D-printed extensions for her faucets (2b) so they open by push not by grab-rotating. Receiving a snapshot of the bathroom, AccessLens recommended a toothpaste squeezer (2d) that alleviates the strain.

Scenario #3: Safety-Proofing Home-office. Arjun, a single individual is preparing to host a home party at his home studio with a safe environment for families with young children. He reviewed the home using AccessLens, identifying potential risks that we never were aware of. Recommendations ranged from switch covers (3a) to prevent sink grinder accidents, child safety stove knobs (3b), machine button covers (3f), and a cord hanger (3c) to prevent the blind cords hazard. He also prepared several 3D-printed bumpers (3e) to be attached to sharp edges, and drawer locks (1f), which can be also useful for Julie. Seemingly innocuous office chair wheels (3d) were also covered by AccessLens.

Scenario #4: Caring for the Family. Mia is a devoted daughter and caretaker of her elderly mother, who is increasingly lacking mobility capacity. AccessLens suggested specialized tools designed to facilitate daily routines: a sock aid (4b) to help avoid bending, a button hook (4c) to simplify fastening shirts, and an extended shoe horn (4d). Additionally, AccessLens recommended a switch extension (4a), allowing her mother to operate it easily without precise hand movements or while on home medical equipment. Mia also used drawer labels (4e) for her bath products for easy identification using larger texts. Enabling greater comfort and independence in managing daily tasks, Mia also found those are overall accessible for her young child.

C   DETECTOR PERFORMANCE

C.1   Evaluation Metrics.

The literature of object detection commonly uses the standard metric of mean Average Precision (mAP) at interaction-over-union (IoU) thresholds ranging from 0.5 to 0.95, with a step size 0.05 [39]. We use mAP as the primary metric. Following other prior works [2, 17, 47], we also report performance with respect to the metrics of AP_{50} and AP_{75} [23], meaning the Average Precision (AP) at IoU threshold 0.5 and 0.75, respectively.

C.2   Training a detector with AccessDB

AccessLens supports detection for all object classes in the 3D assistive augmentation dictionary and ICs. Although any detector structure can be chosen, we utilized two different state-of-the-art methods, RetinaNet [37] for ICs with training on AccessDB, and GroundingDINO [40] for zero-shot detection without training for more common classes (e.g., sofa, table, cup, etc.) Specifically, we trained RetinaNet [37] with ResNet-50-FPN backbone with 3x LR schedule, implemented by detectron2 [75]. In training, we employed COCO pretrained weights retrieved from Model Zoo of detectron2. As Figure 8 illustrates, AccessDB contains 21 inaccessibility classes, and one extra class, unidentifiable instances due to their extremely small size to identify with human eyes. For training and validation of the detector, we randomly split the dataset, 85% training and 15% validation (2,029 and 359 images for training and validation, respectively). We used the AccessReal dataset for testing (42 images) to understand and compare how the detector works on AccessDB and more high-resolution images in AccessReal. For the 'unidentifiable' class, we still included it as an individual class in training but did not use it for evaluation. This is because, 'unidentifiable' objects are still in the 6 categories of our interests, so those might have overlapping visual features with other inaccessibility classes that

id	inaccessibility class	AccessDB	AccessReal
1	button_panel_push_buttons	13.91	11.43
2	button_panel_turn_handle	26.48	7.72
3	electric_outlet	29.94	16.65
4	faucet_faucet_only	21.36	4.90
5	faucet_handle_lever	29.92	12.85
6	faucet_pull_tiny_knob	38.52	n/a
7	faucet_rotate_cross	34.84	n/a
8	faucet_rotate_knob	32.92	n/a
9	handle_bar_large	13.04	5.7
10	handle_bar_small	16.78	10.37
11	handle_cup_handle	3.21	0.04
12	handle_drop_pull	27.99	n/a
13	handle_flush_pull	0.80	n/a
14	handle_lever	12.38	15.71
15	handle_pull	9.40	1.03
16	knob_rotate_round	29.08	23.01
17	knob_static	16.34	2.26
18	switch_rocker_multi	14.31	23.50
19	switch_rocker_single	1.53	2.02
20	switch_toggle_multi	31.36	64.21
21	switch_toggle_single	10.52	36.20
	average	18.85	14.86

Table 4: Breakdown results of our inaccessible-object detector on AccessDB validation set and AccessReal. Performance is measured by AP for each inaccessibility class. AP metrics on AccessDB are generally higher than AccessReal, showing a reasonable domain gap. Yet, on some inaccessibility classes such as switch_toggle_single and switch_toggle_multi, AP metrics on AccessReal are higher, presumably because images of AccessReal are higher in resolution that these small inaccessible objects are clearer and easier to detect than AccessDB images.

C.3 Detector Analysis

Evaluation of the detector on validation and test sets was performed per each epoch. The detector achieved its best performance for the AccessReal dataset after around 51 epochs, yielding an mAP of 18.85 for the validation set and 14.86 for the test set. Additional performance metrics are provided in Table 5. AccessDB validation set showed the best mAP (19.86) at epoch 61, but after 51 epochs the detector started overfitting to AccessDB, resulting in the lower mAP (13.36) for AccessReal. Even though AccessDB and AccessReal both contain real-world indoor images, we could still see the domain gap between the two as the detector shows about 4 less mAP. We attribute this performance difference, in part, to the significantly higher resolution of images in AccessReal, which poses a challenge for a detector primarily trained on smaller images. Furthermore, AccessDB inherently exhibits a long-tailed distribution in terms of class counts (Detailed breakdown of the number of classes is described in Table 3). This distribution presents an additional challenge to the detector, particularly when recognizing classes with a relatively small number of objects, which may not provide sufficient data for the model to learn distinctive visual features. Despite the challenges, visual results created by our detector (Figure 13) showcase its ability to perform well on high-resolution indoor images. In the zoomed regions of Figure 13 (second and fourth images), results show that the detector successfully recognized our interested objects, including knob_rotate_round, faucet_handle_lever, electric_outlet, and handle_bar_small. Table 4 provides a breakdown of mAP for each inaccessibility class. The average mAP indicates that, as a whole, the detector performs better on AccessDB compared to AccessReal. However, it's worth noting that the detector exhibits superior performance on AccessReal for certain classes, such as switch_toggle_multi, switch_toggle_single, and handle_lever. We hypothesize that for these classes, AccessReal may offer clearer object representations or exhibit fewer visual variations, possibly due to its smaller sample size, thereby contributing to improved detection accuracy.

	mAP	AP_{50}	AP_{75}
AccessDB	18.85	33.41	19.03
AccessReal	14.86	28.24	11.55

Table 5: We evaluate our inaccessible-object detector (based on the RetinaNet architecture [37]) on the validation set of AccessDB, and the AccessReal (as the testing set). Quantitative results show a clear domain gap between the two datasets; visual results in Figure 13 demonstrate that our detector (trained on AccessDB's training set) can detect inaccessible objects quite well in AccessReal, representing modern indoor scenes.

the human eye could not capture due to the blurry images. By treating it as one class in training a detector, we can avoid unwanted penalizing of the other classes' correct predictions.