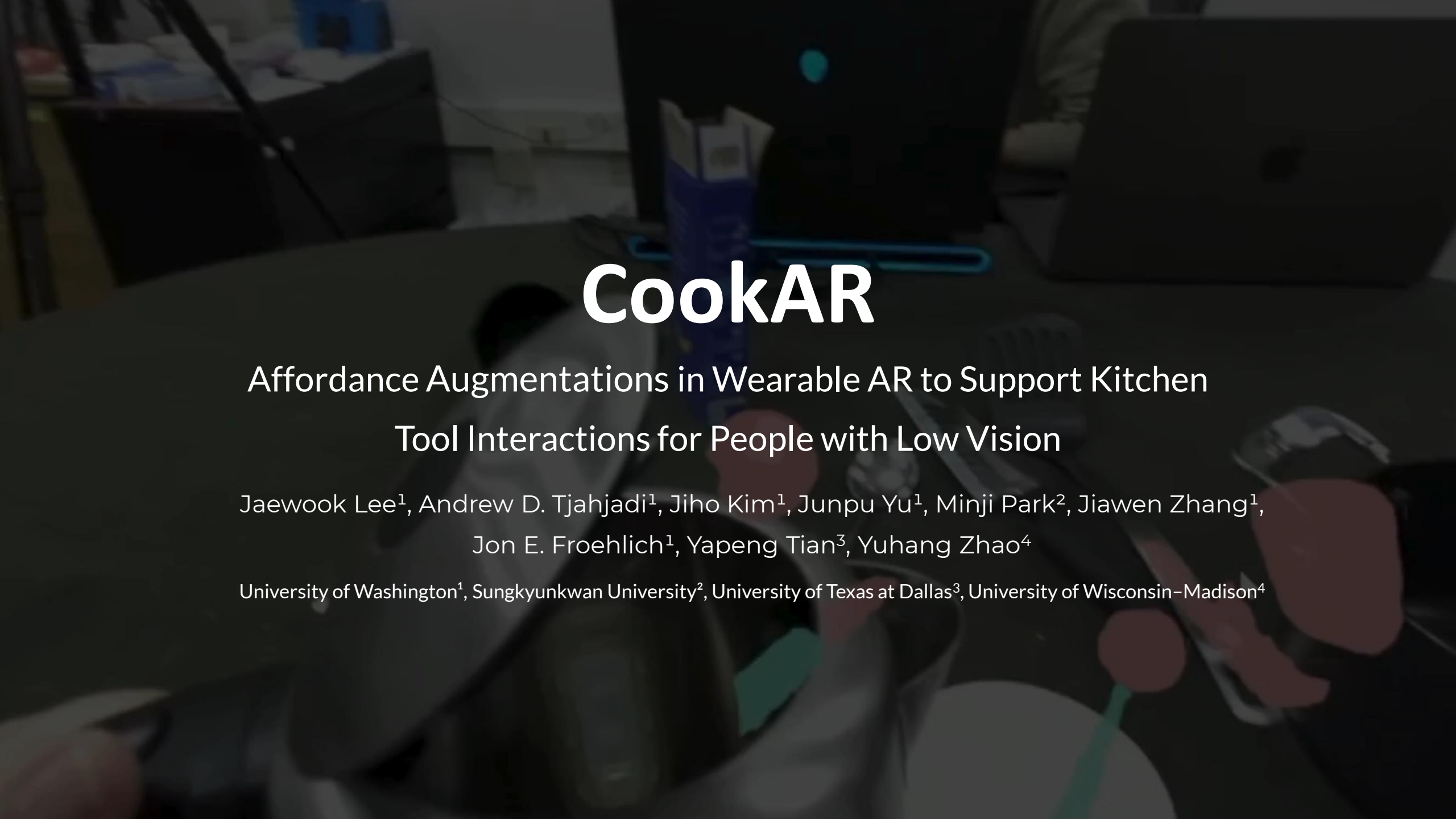


CookAR

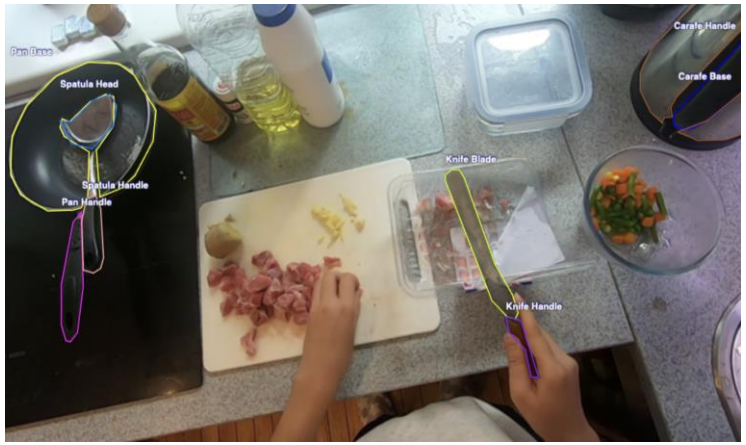
Affordance Augmentations in Wearable AR to Support Kitchen Tool Interactions for People with Low Vision

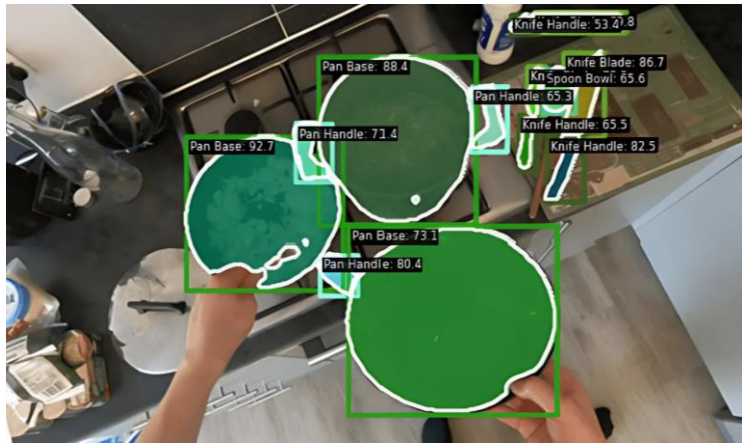
Jaewook Lee[1], Andrew D. Tjahjadi[1], Jiho Kim[1], Junpu Yu[1], Minji Park[2], Jiawen Zhang[1], Jon E. Froehlich[1], Yapeng Tian[3], Yuhang Zhao[4]

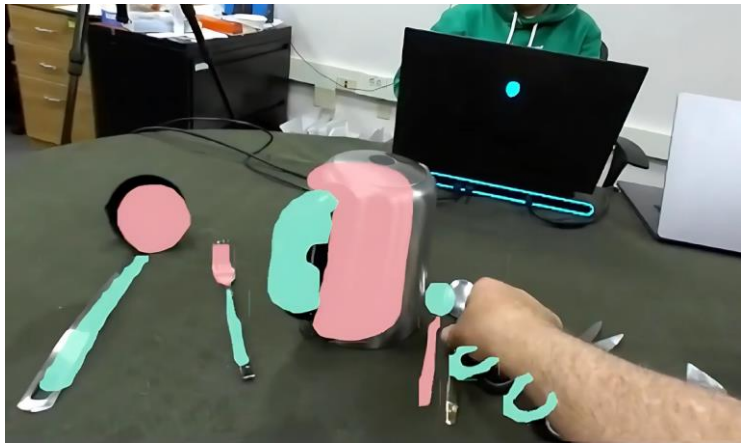
University of Washington[1], Sungkyunkwan University[2], University of Texas at Dallas[3], University of Wisconsin–Madison[4]

We introduce ***CookAR***, a wearable AR prototype that recognizes and augments cooking tool affordances in near real-time to support LV cooking.

The CookAR System







01 A custom egocentric kitchen tool affordance dataset of 10,152 images.

02 A fine-tuned affordance segmentation model.

03 A wearable AR research prototype capable of overlaying visual augmentations.

Data Collection

We start with the *Epic Kitchens* dataset, an egocentric video dataset of sighted people cooking.

Damen et al. The EPIC-KITCHENS Dataset: Collection, Challenges and Baselines

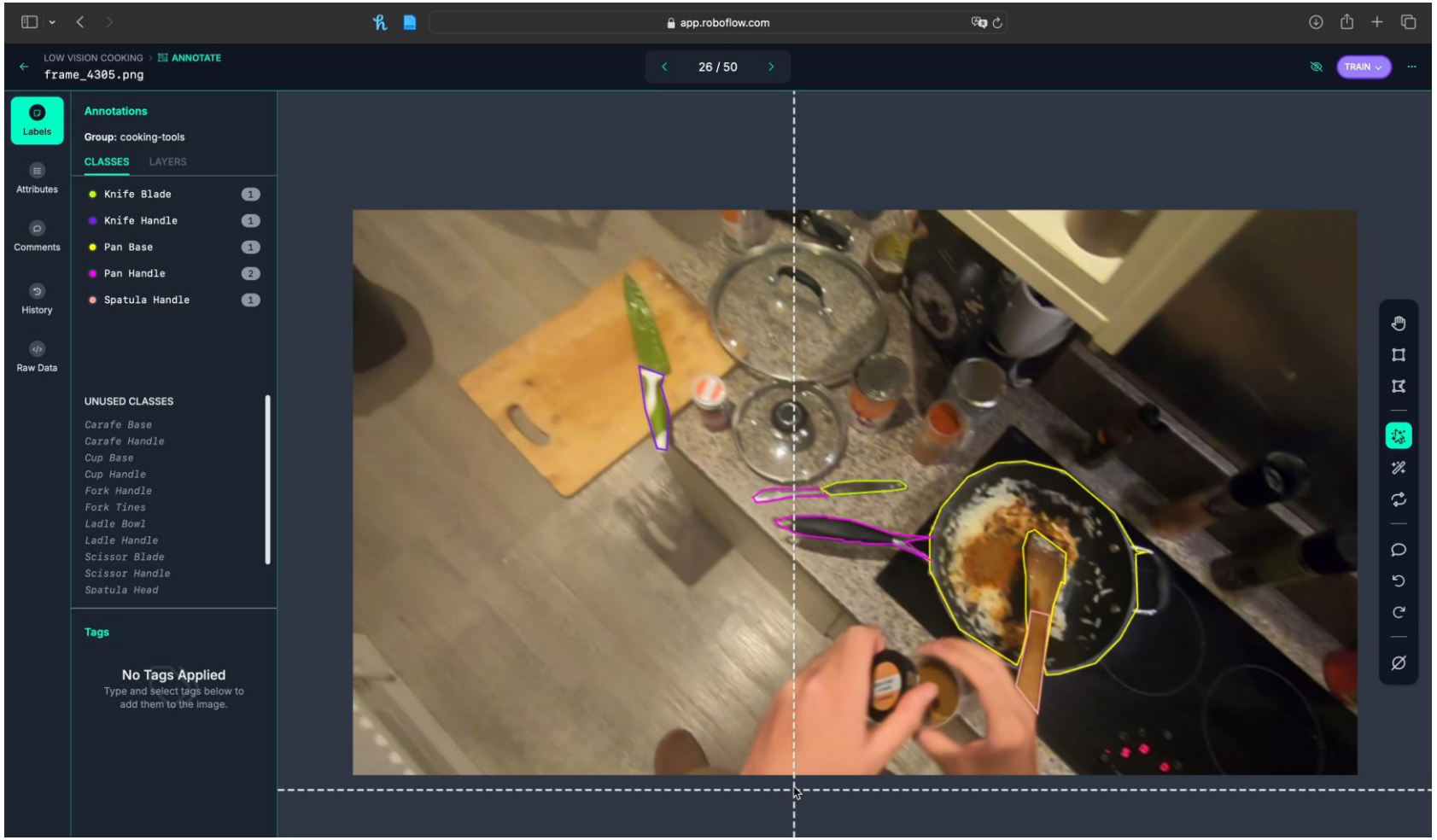
Data Collection

We first use *YOLOv8* to find key frames in the Epic Kitchens dataset, followed by manual filtering.

YOLOv8 inferencing on Damen et al.'s video wall

Data Annotation

We then used *Roboflow* and the *Segment Anything* model (*SAM*) to manually label selected frames.



Data Annotation

We labeled 18 distinct affordance classes of kitchen tools commonly used by BLV individuals.

Knife Blade & Handle

Spoon Bowl & Handle

Fork Tines & Handle

Scissor Blade & Handle

Ladle Bowl & Handle

Spatula Head & Handle

Pan Base & Handle

Cup Base & Handle

Carafe Base & Handle

Model Fine-Tuning

We fine-tuned the *RTMDet-ins-l* model, the current state-of-the-art in real-time instance segmentation.

Undergoing fine-tuning on a dataset with affordance annotations allows it to mimic an affordance segmentation model's capabilities.

Example inferencing result of our custom RTMDet-ins-l-cook model.

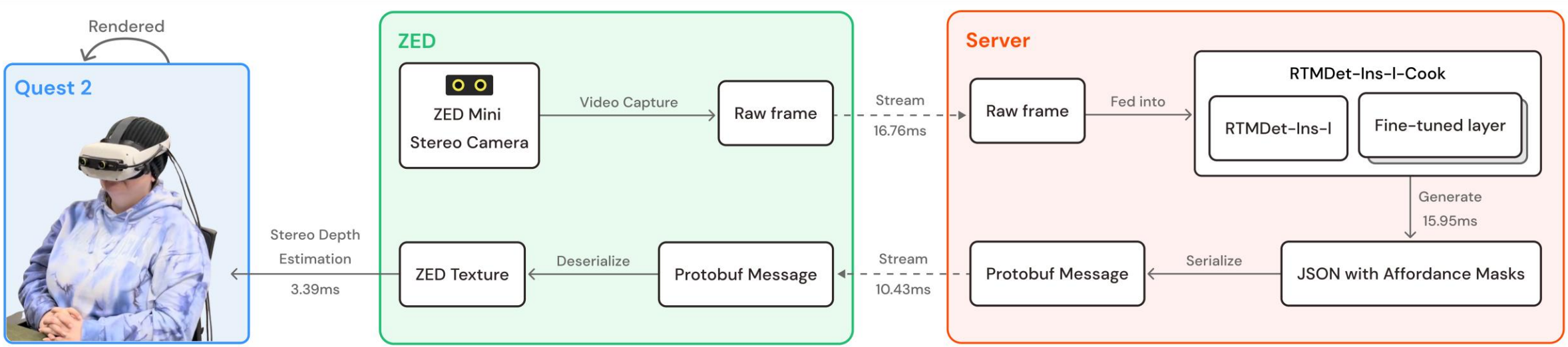
0.605 FPS
Pan Base: 94.9
Pan Handle: 32.6
Pan Handle: 72.5
Spoon Bowl: 67.1
Pan Handle: 79.8
Fork Tines: 58.4
Fork Handle: 34.6
stir chicken
Pan Base: 35.3
Pan Base: 76.1
Carafe Base: 68.1
Carafe Handle: 35.0
press down on aeropress
Pan Base: 87.0
Pan Handle: 51.4
Spoon Bowl: 36.2
Spatula Handle: 30.5
soak pan
Cup Base: 56.0
Cup Base: 31.7
Cup Base: 34.9
Cup Base: 33.6
Cup Base: 67.4
Cup Base: 30.2
put down sponge
Knife Blade: 43.7
Fork Tines: 31.0
Spoon Bowl: 32.3
Spoon Handle: 40.1
put scotch egg on plate
Pan Base: 84.5
wipe down counter
Knife Blade: 36.6
Knife Handle: 45.1
Cup Base: 86.6
Cup Handle: 70.2
stretch dough
Pan Base: 59.1
Pan Base: 63.5
place packet of cumin seeds on shelf

AR Prototype

We built a custom stereo video see-through AR system by combining the ZED Mini stereo camera with an Oculus Quest 2 headset.

The ZED Mini stereo camera enables CookAR to perform binocular depth estimation.

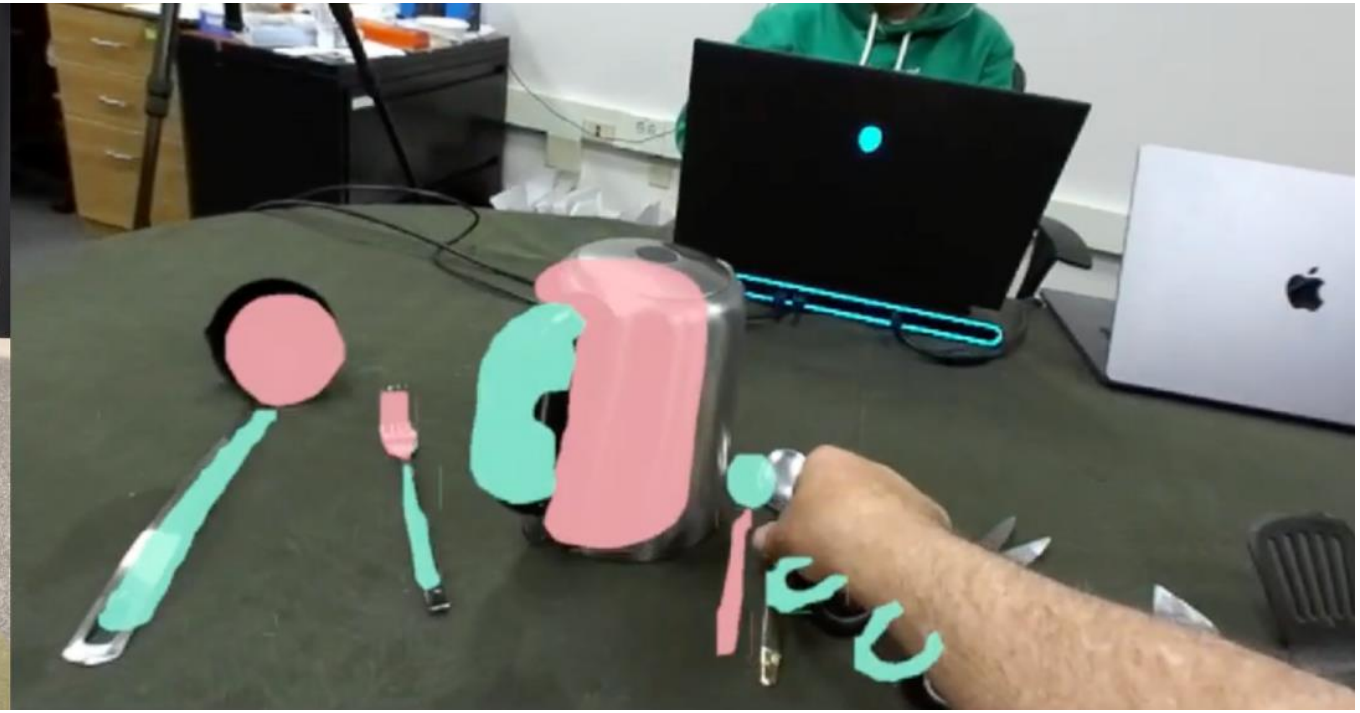
We stream raw video frames from ZED to an external server for heavy CV.



AR Prototype

The CookAR AR prototype displays affordance augmentations:

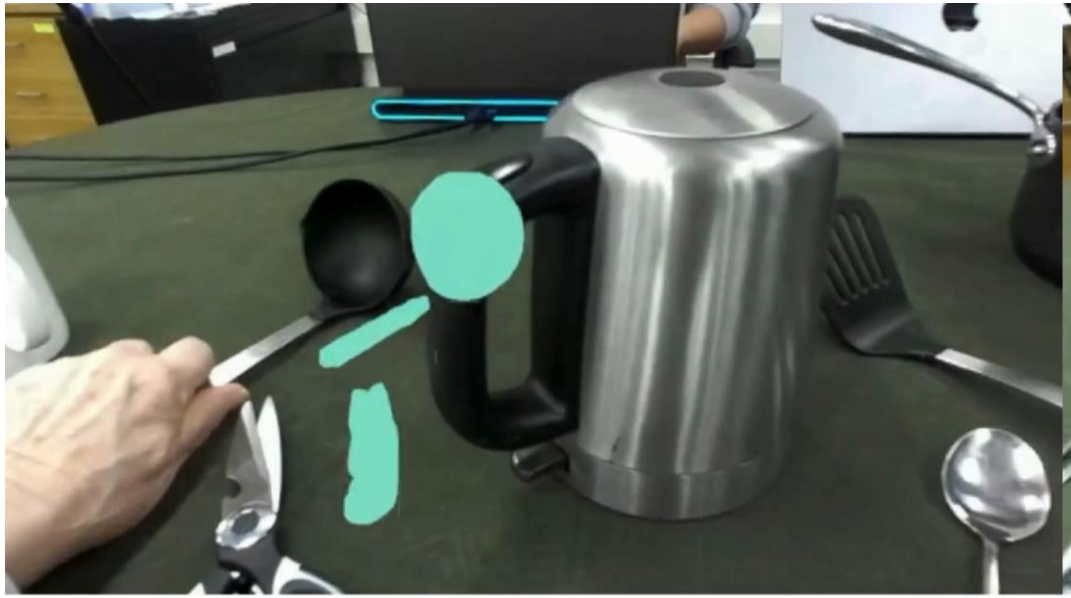
- **Green** (hexcode #3BE8B0) indicates a graspable area.
- **Red** (#FC626A) indicates a risky area.



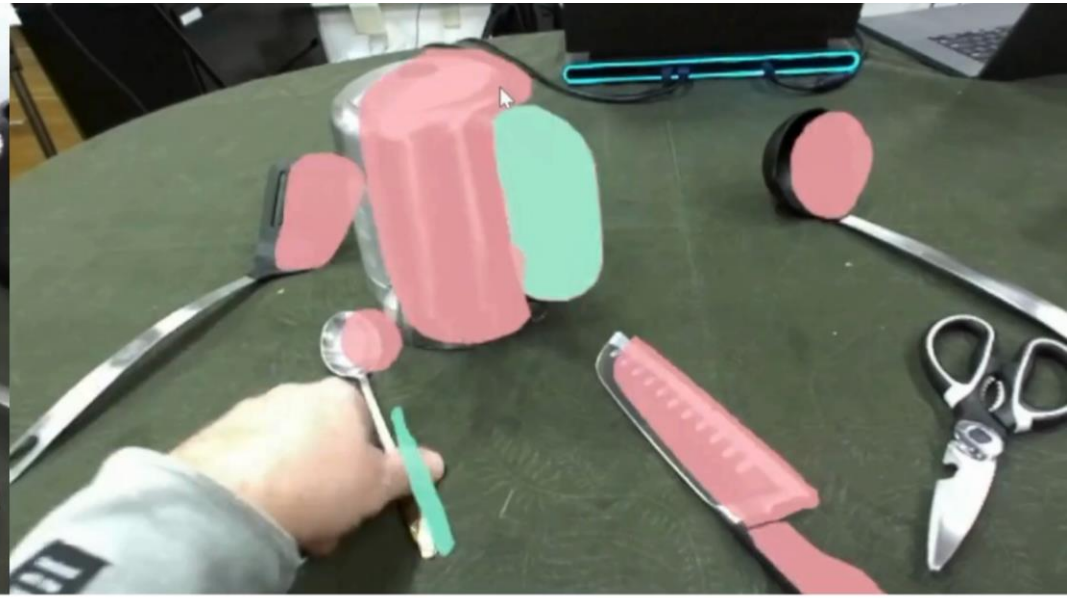
Life is good

User Study

We conducted a three-part lab study with 10 low vision participants.



Whole Object Augmentations



CookAR with Affordance Augmentations

Part 1 - comparing three conditions in a tool grabbing task.

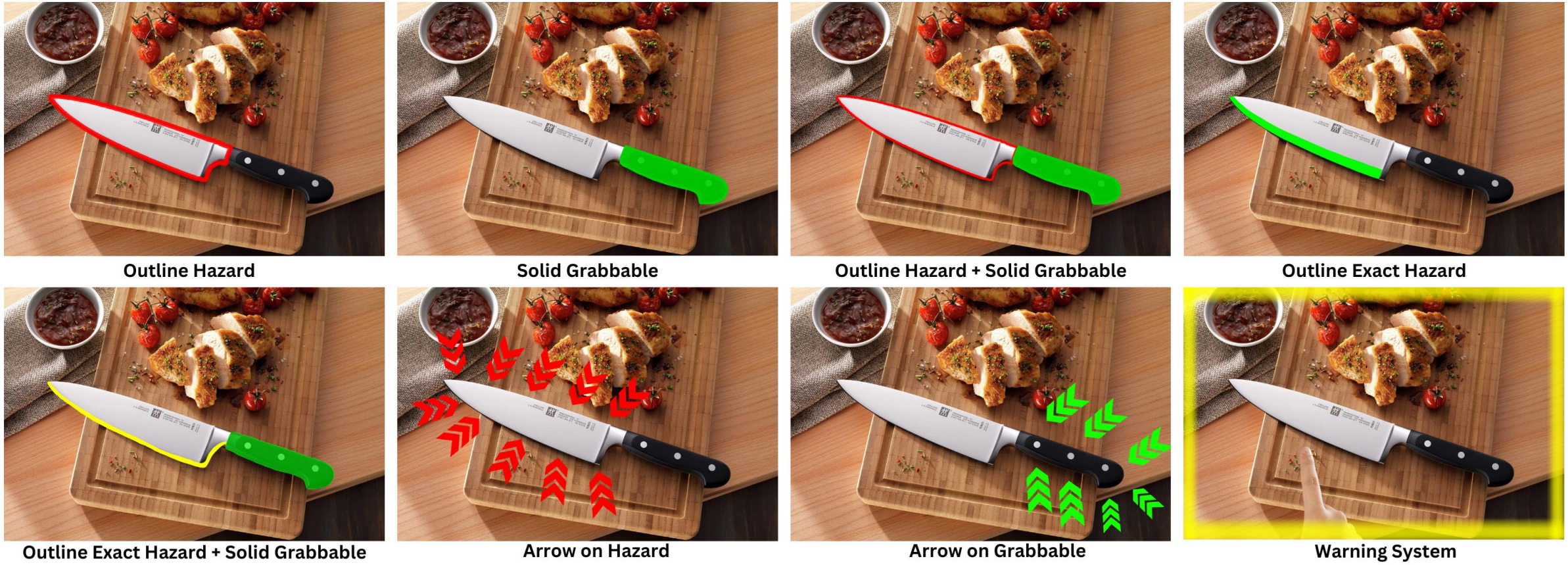
1. Real world baseline (i.e., without CookAR),
2. Augmentation baseline (i.e., whole object augmentations),
3. CookAR (i.e., affordance augmentations)

User Study

Part 2 - cooking macaroni and cheese using CookAR with affordance augmentations.

User Study

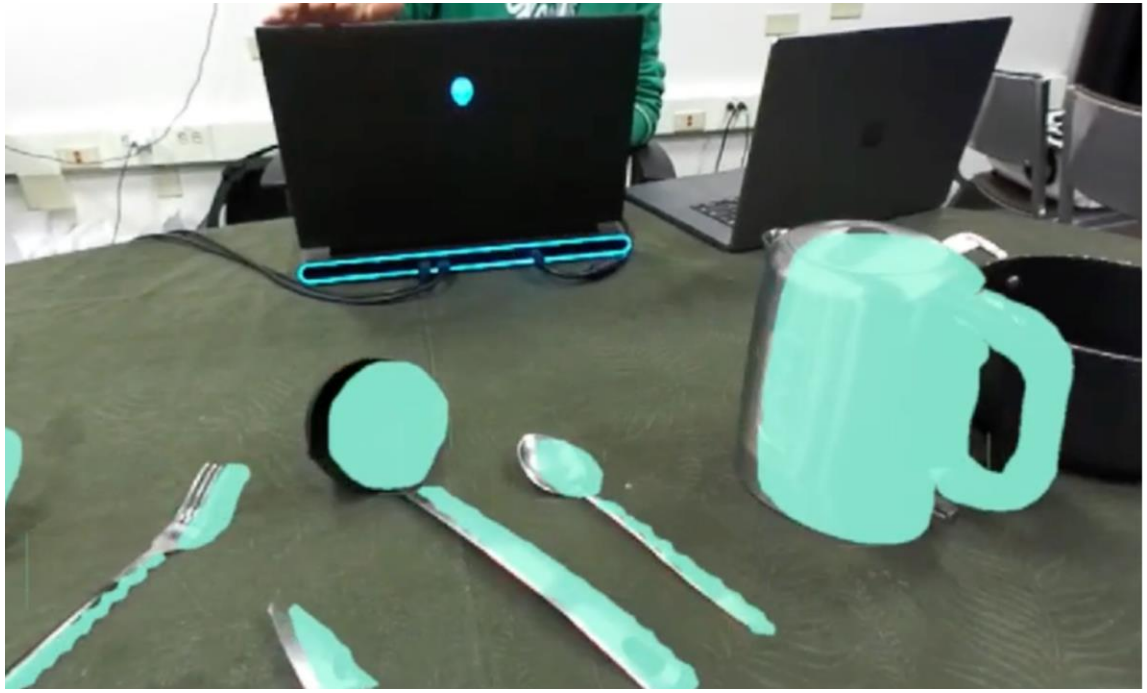
Part 3 – brainstorming and co-designing future designs and applications of CookAR.

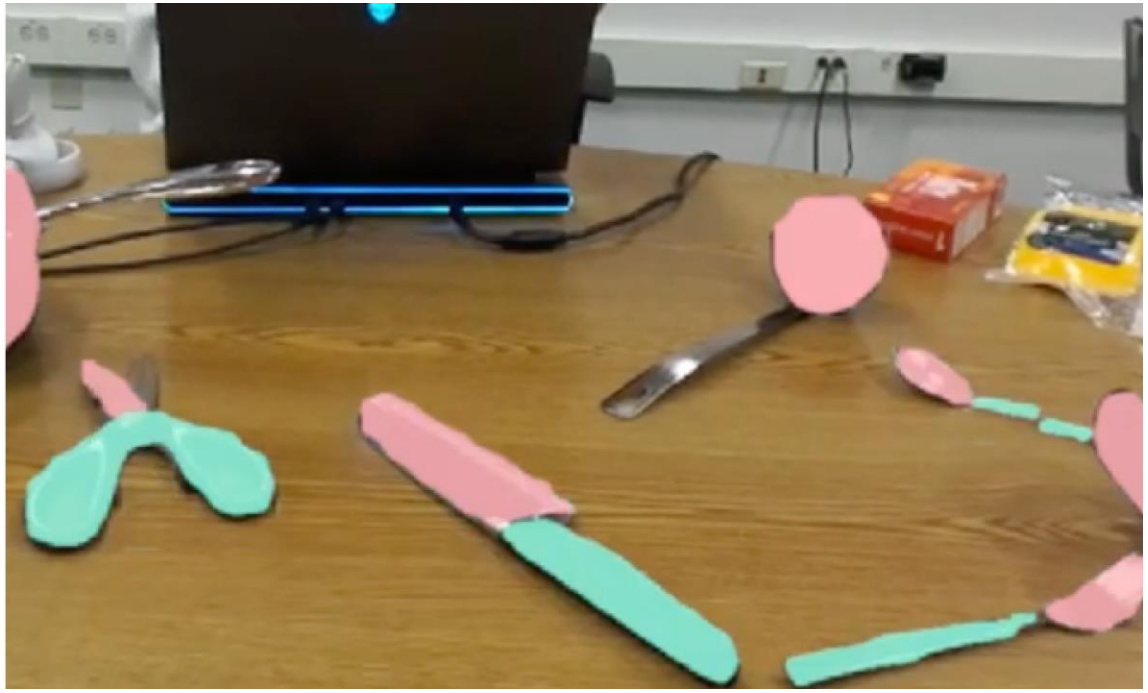


Affordance vs. Whole Object Augmentations

CookAR should render augmentations tailored to users' intent and reduce distraction.

- For complex objects and hand-object manipulations, **affordance augmentations** are better.
  - Gardening, playing board games, interacting with appliances, etc.
- For general visual perception, **whole object augmentations** are better.
  - Avoiding obstacles, locating an object, etc.

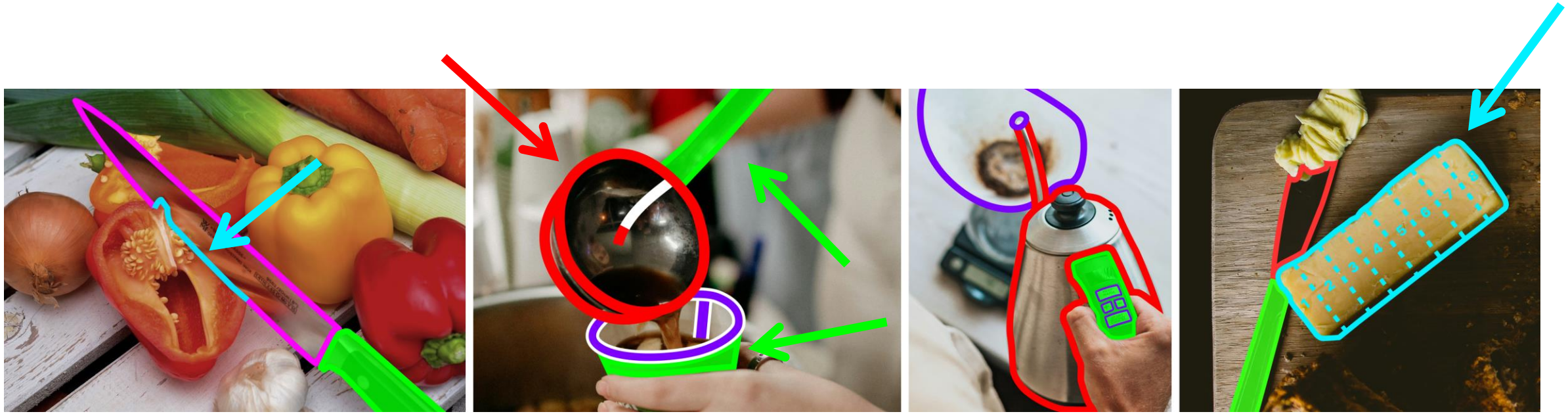




Desired Augmentation Designs

Participants desire the following augmentation designs:

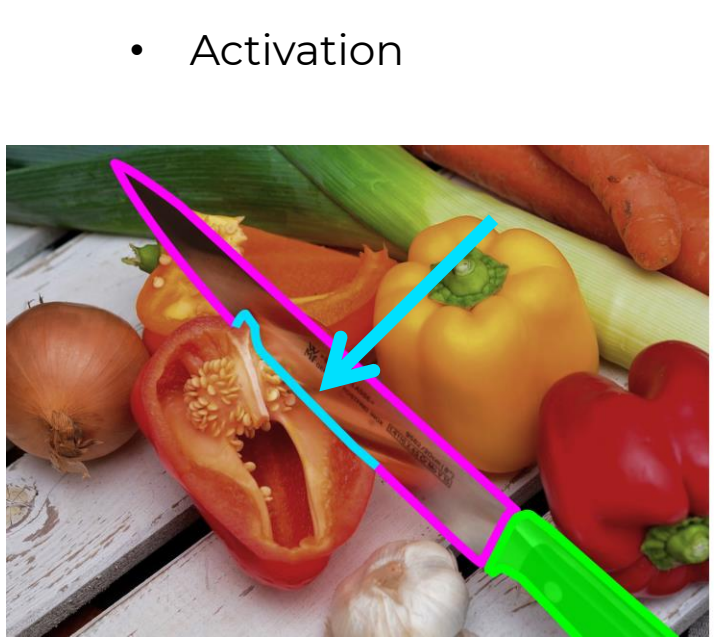
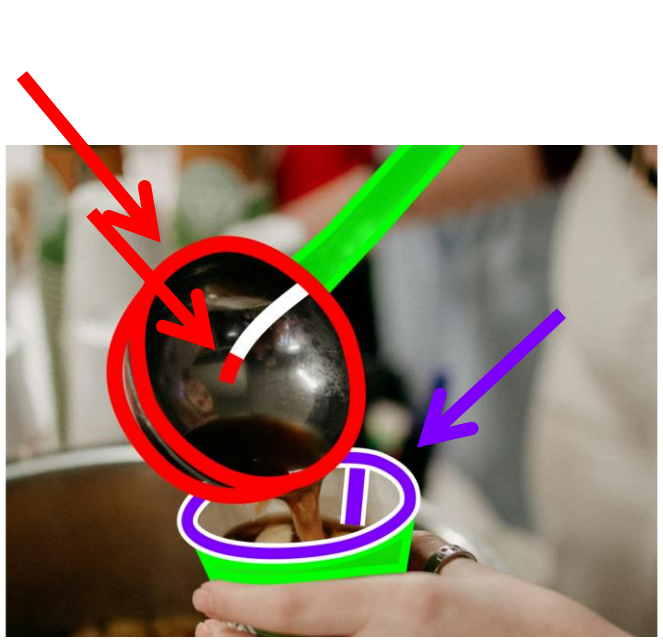
- Combine solid and outline augmentations
- Enhance color contrast
- Auditory feedback
- Personalization

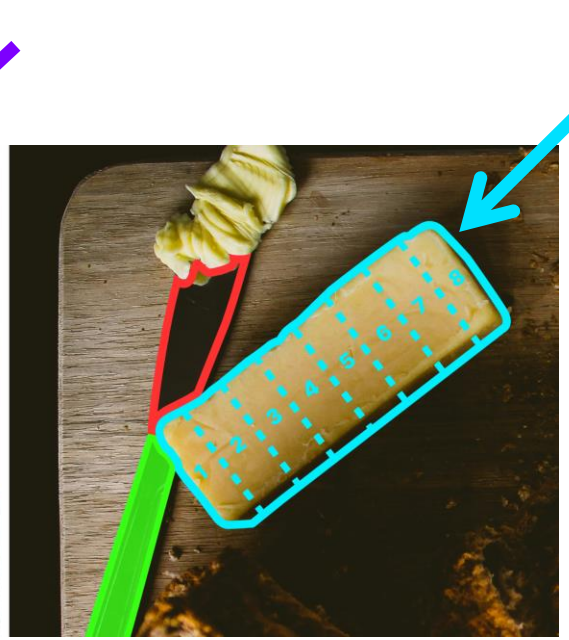


Additional Tool Affordances

Beyond grabbable and hazardous areas, participants identified five additional tool affordances:

- Entry
- Exit
- Containment
- Intersection
- Activation



CookAR

Affordance Augmentations in Wearable AR to Support Kitchen Tool Interactions for People with Low Vision

Jaewook Lee [1,2], Andrew D Tjahjadi [1,2], Jiho Kim [1,2], Junpu Yu [1,2], Minji Park[1,3], Jiawen Zhang [1,2],
Jon E. Froehlich [1,2], Yapeng Tian[4], Yuhang Zhao[5]

MAKEABILITY LAB
UNIVERSITY of WASHINGTON
SUNGKYUNKWAN UNIVERSITY(SKKU)
THE UNIVERSITY OF TEXAS AT DALLAS
THE UNIVERSITY of WISCONSIN MADISON

Dataset & Model

Thanks for listening!

Let’s Connect!

Paper

Email jaewook4@cs.washington.edu

Website https://jaewook-lee.com

Twitter https://twitter.com/jaewook_jae