

Mobile and Ubiquitous Computing – Midterm Presentation

# NutriSeeON

## virtual eyes to consume healthy groceries

**Team 3**

Baek Eunsu

Kim Juhee

Park Seokhyeon



# Project Idea & Scope

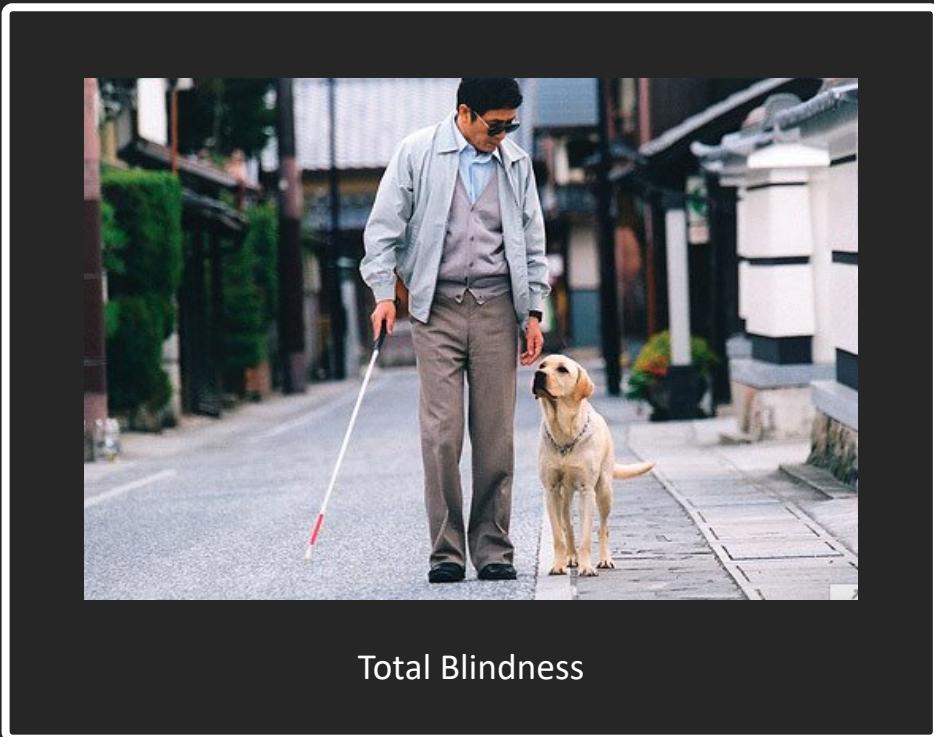
# Project Idea

- Shoppers buy groceries considering their health concerns.
  - Weight control, disease (diabetes, high blood pressure), allergies, ...
- Checking **Nutrition Facts Table** is one of the common ways to choose the right food.



# Target User

Is it possible for blind people to purchase products by checking the nutrition table?



Total Blindness



Healthy Shopping

# Solution Summary

NutriSeeON will act as the eye of blind users, helping them to shop healthy groceries.

- Real-time Nutrition Facts Extraction by
  - Mobile Camera
  - Tactile / Auditory Feedback
  - Deep learning based Image processing



Find out nutritional facts label  
as if they had a full vison



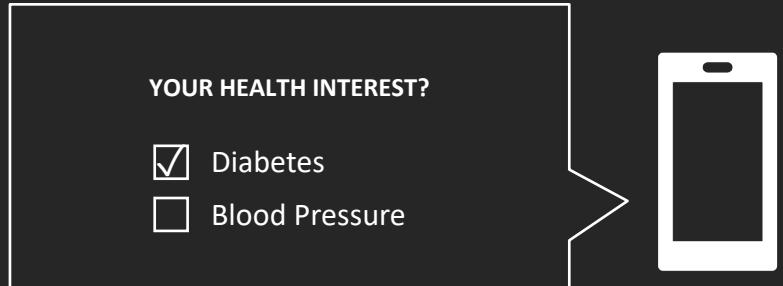
Inform nutritional information to be aware  
of based on their health concern

# Usage Scenario

# Usage Scenario

## Step 1

Set user health concerns.  
(e.g., Diabetes, high blood pressure, weight loss, ...)



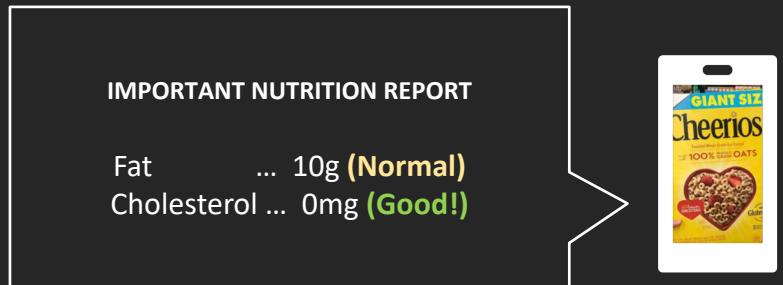
## Step 2

Pick up food and take a picture of it.  
It helps user to take photos of the food.



## Step 3

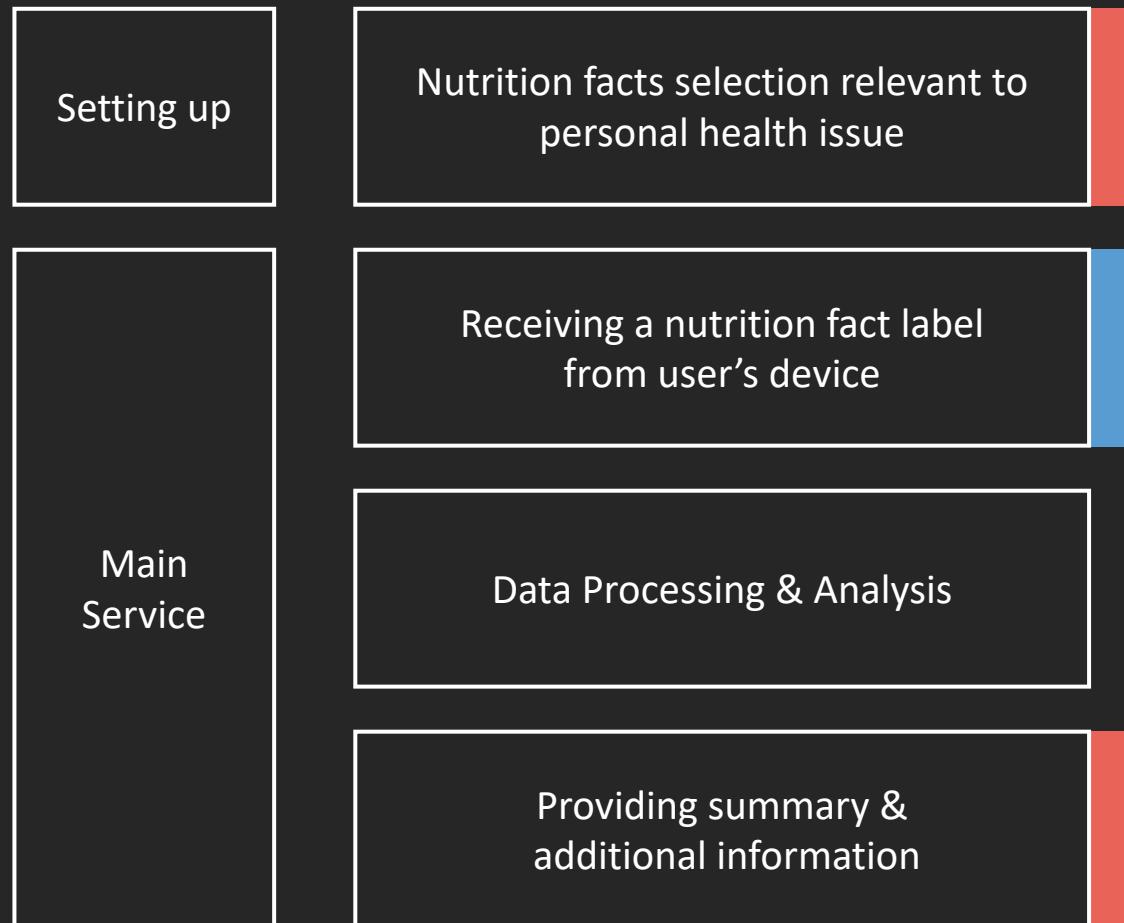
It collects nutritional information from photos and informs the information that suits the user's interest.



# System Overview

# System Overview & Main Functions

## System Overview



## Main Functions



Tactile/Auditory based  
Nutrition Facts  
Recognition Guide Module



Auditory support  
for interaction

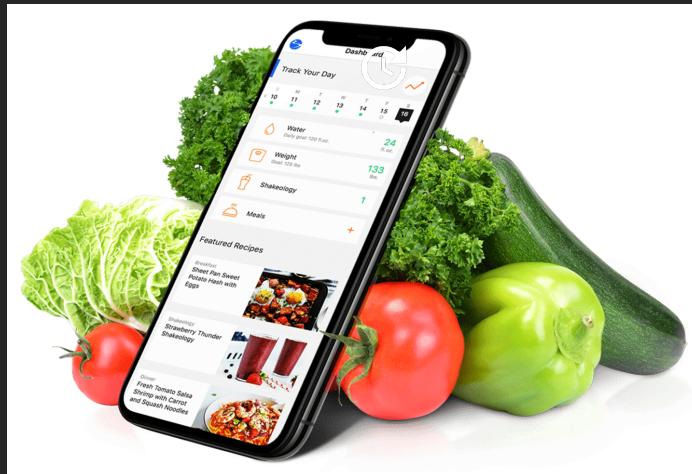


Vibration board with  
directional expression

# Technical Challenges & Solutions

# Technical Challenges

1. Recognizing target object in wild environment
2. Providing intuitive feedback for target users to find nutrition facts label
3. Gathering nutritional information from photos taken in real-world
4. Effectively delivering nutritional information to visually impaired user



# Solution ideas with challenges

## 1. Recognizing target object in wild environment (e.g., grocery shop)

- Original plan : Using public model<sup>[1]</sup> and public dataset<sup>[2]</sup>
- Hand segmentation and object localization
  - give feedback **to locate object on the proper position**



Hand Segmentation



Object Localization



[1] Lee, Kyungjun, and Hernisa Kacorri. "Hands Holding Clues for Object Recognition in Teachable Machines." *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 2019.

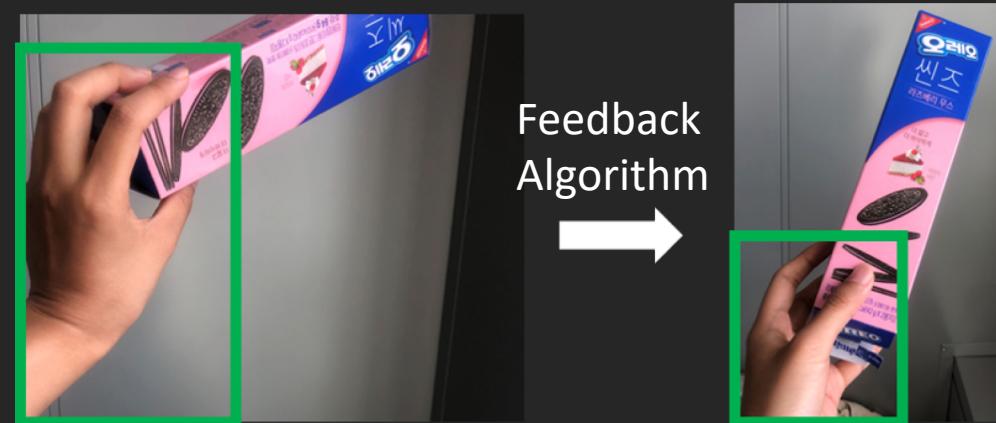
[2] Tego

# Solution ideas with challenges

## Challenge 1: Difficulties in *Hand-hold Object Localization*

- No useful public model and no labeled public dataset (TEgO: lacks labels)
- Resource constraints : Lack of GPU, not much time to create datasets
- Our solution: Using the public *Hand Detection model*
  - hand localization from the public model
  - From the identified hand location, give feedback **to locate hand on the proper position**

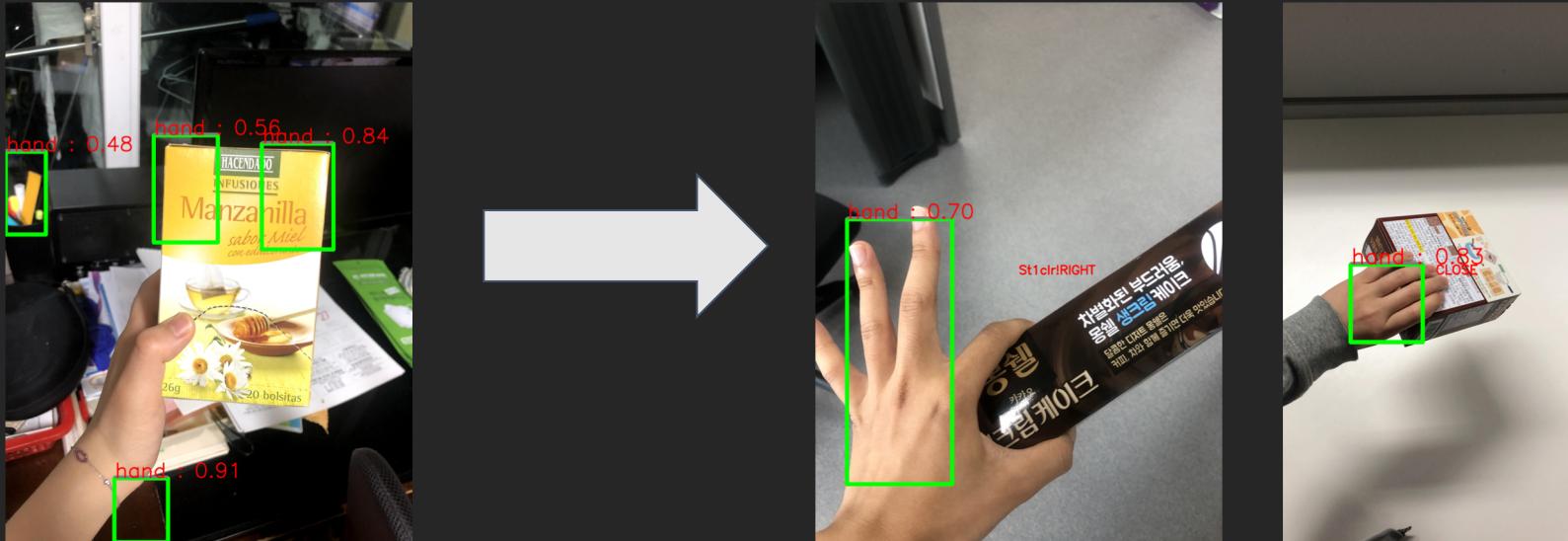
Hand Detection



# Solution ideas with challenges

## Challenge 2: Poor accuracy of the hand detection model

- Our solution:
  - Obtain stable hand detection results by post processing
  - Provide guideline for grip to easily identify hand
    - with the back of your left hand facing up,  
Ok grip (grip object with tweezer part) / All finger grip
  - Constrain the environment setting (Clean background)



# Solution ideas for challenges

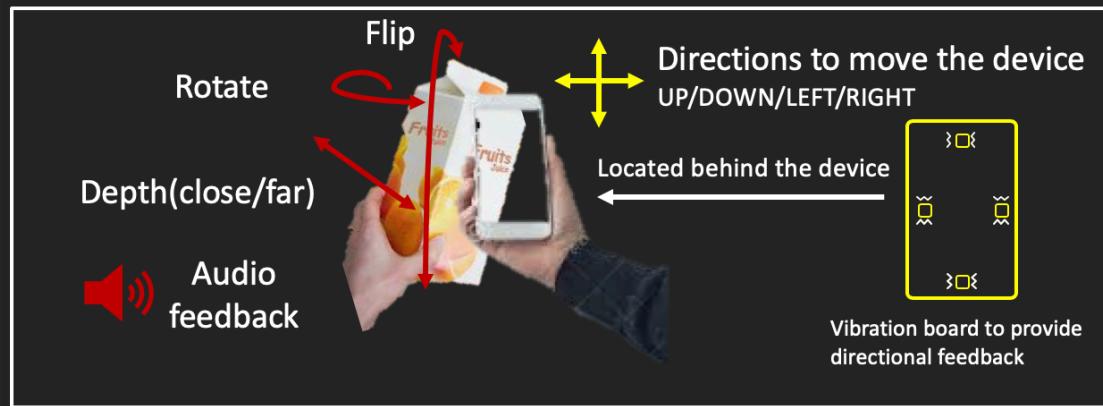
## 2. Intuitive feedback system for target users to find nutrition facts label

- Divide the feedback into 2 types → *DONE*
  - 1) Feedback on the *object* (left hand)
  - 2) Feedback on the location of the *camera* (right hand)
- Deliver the tactile & auditory feedback directly to each hand
  - 1) Tactile feedback wt. a wearable device (right hand) → *Audio feedback only*
  - 2) Vibration board to inform the direction to move the camera (left hand) → *DONE*



Tactile Flow [1]

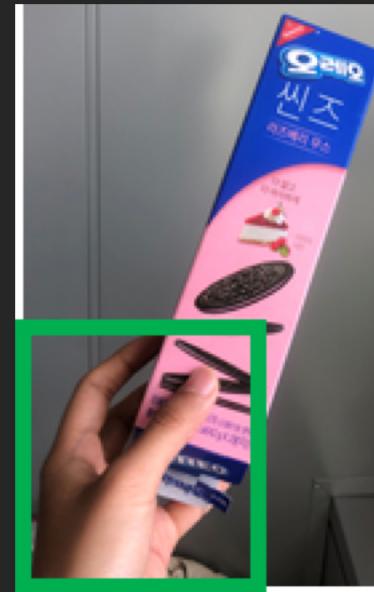
# Feedback Algorithms



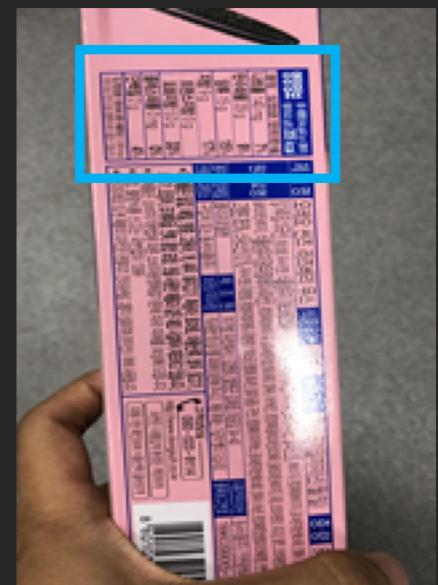
detect hand  
& size control  
with close/far  
feedback



locate the  
left hand  
at the  
*bottom left*  
of the screen

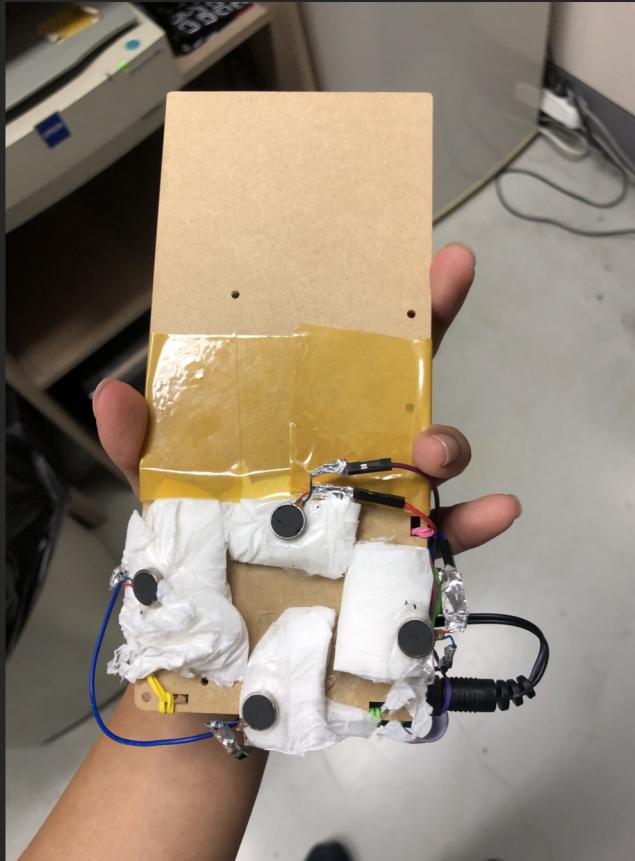
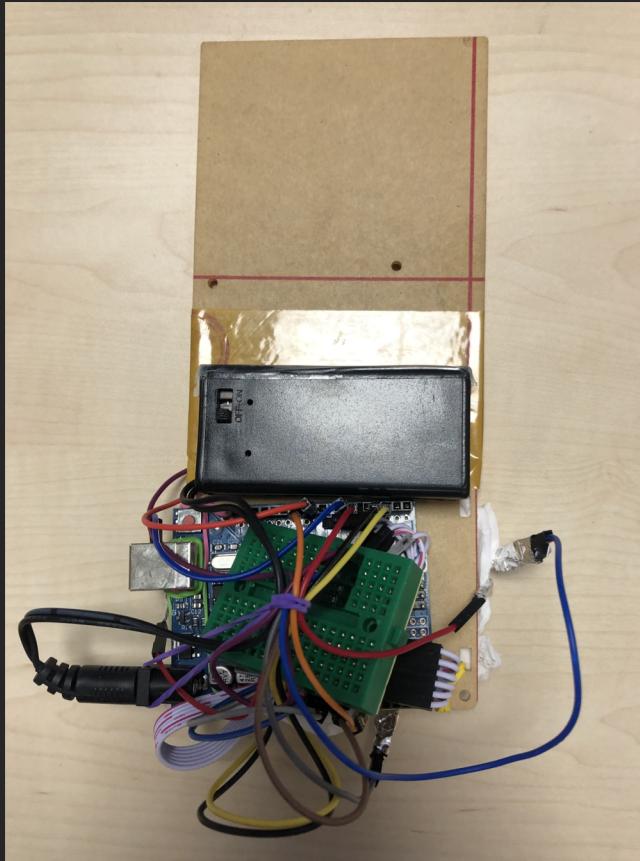


rotate and  
flip until  
it finds the  
nutrition table



# Feedback Module Deliver

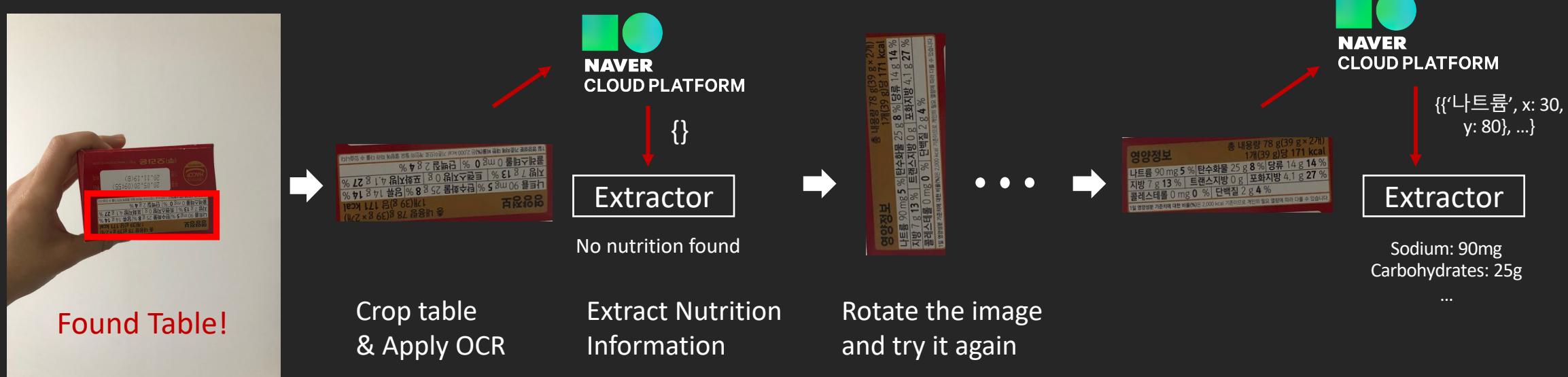
1. Changed Point : removed a vibration motor from watch module due to hardware issue and time constraints.



# Solution ideas for challenges

## 3. Gathering nutritional information from photos taken in real-world

- First, find a nutrition facts table with a light deep learning model<sup>[1]</sup>
- If table is found, apply Optical Character Recognition (OCR)<sup>[2]</sup> for the table
- Extract nutritional information through the coordinates of the recognized words
- If the result is insufficient, rotate image and try it again



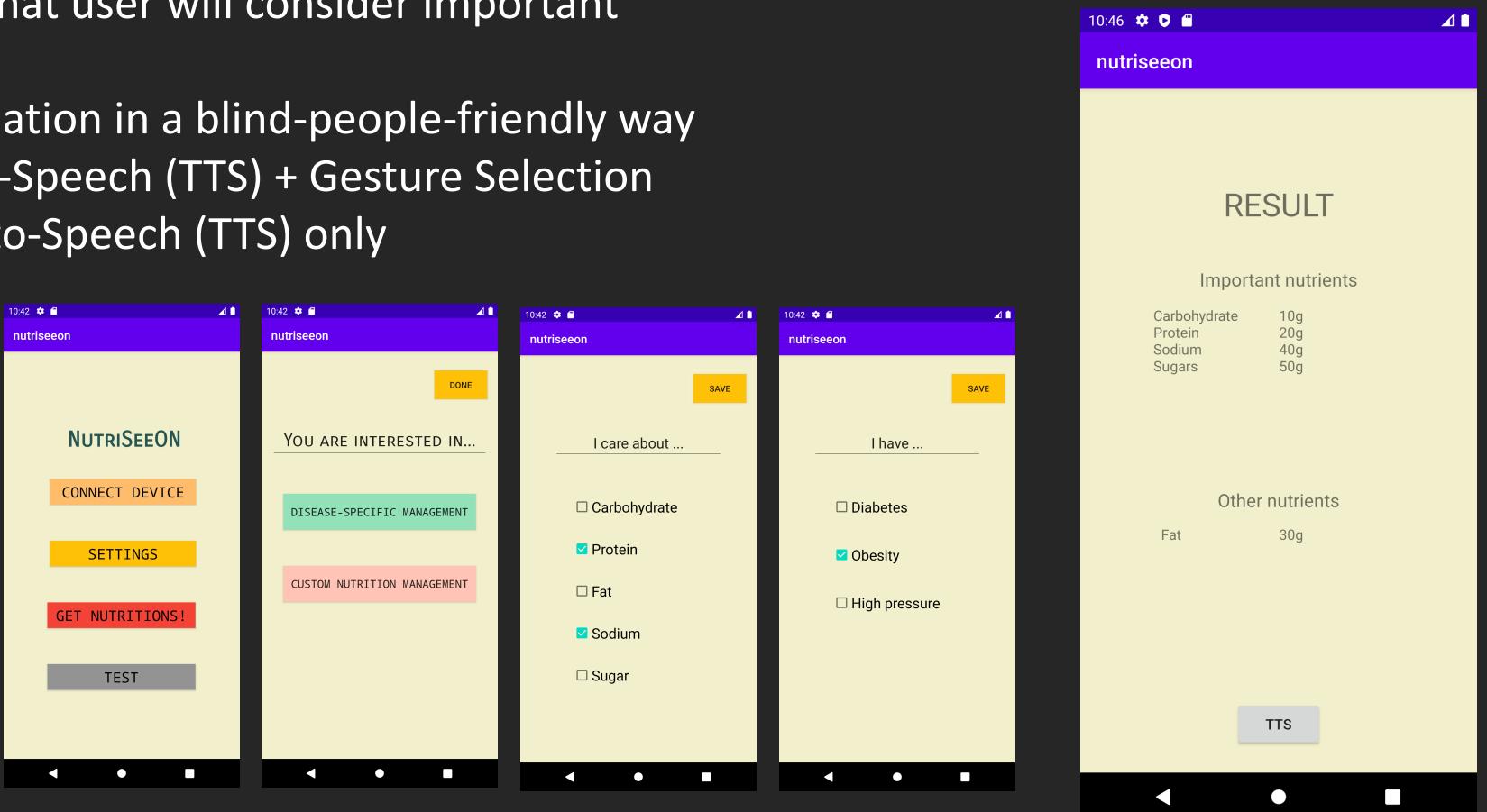
[1] <https://github.com/openfoodfacts/off-nutrition-table-extractor/pull/21>

[2] Naver Clova

# Solution ideas for challenges

## 4. Effectively delivering nutritional information to visually impaired users

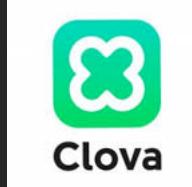
- Provide personalized nutritional information based on user's health concern
  - : Prioritize information that user will consider important
- Deliver nutritional information in a blind-people-friendly way
  - Original idea: Text-to-Speech (TTS) + Gesture Selection
  - Current status: Text-to-Speech (TTS) only



# NutriSeeON Architecture Design



NAVER  
CLOUD PLATFORM



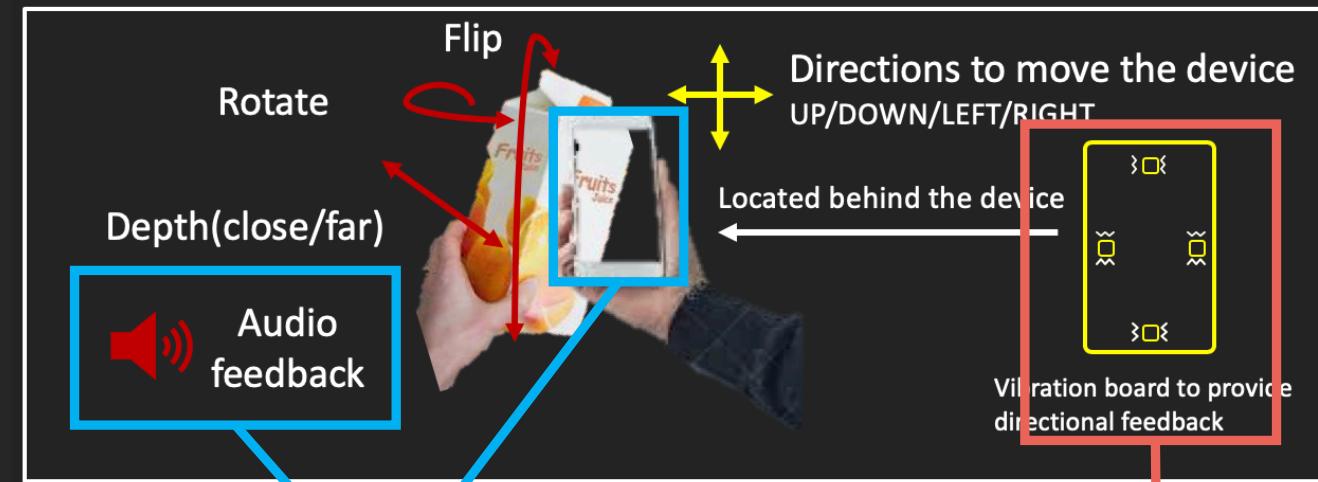
Naver Clova OCR service



NAVER  
CLOUD PLATFORM



Server (CPU) for  
Image processing  
ZEON 5220 2 Core  
RAM 8GB SSD 50GB



1 BLE-available  
Android device  
(tested on Galaxy S10, S9)



Arduino Uno  
SMD



HM-10(BLE 4.0)



9V Battery



x4

4 coin motors

Vibration Board for Camera feedback

# Evaluation

# Final Deliverable & Success Criteria

## Final Deliverable



Mobile Application



Intuitive Feedback Module



Server

Application  
that effectively helps blind people to discover nutritional facts  
and easily get customized information.

## Success Criteria



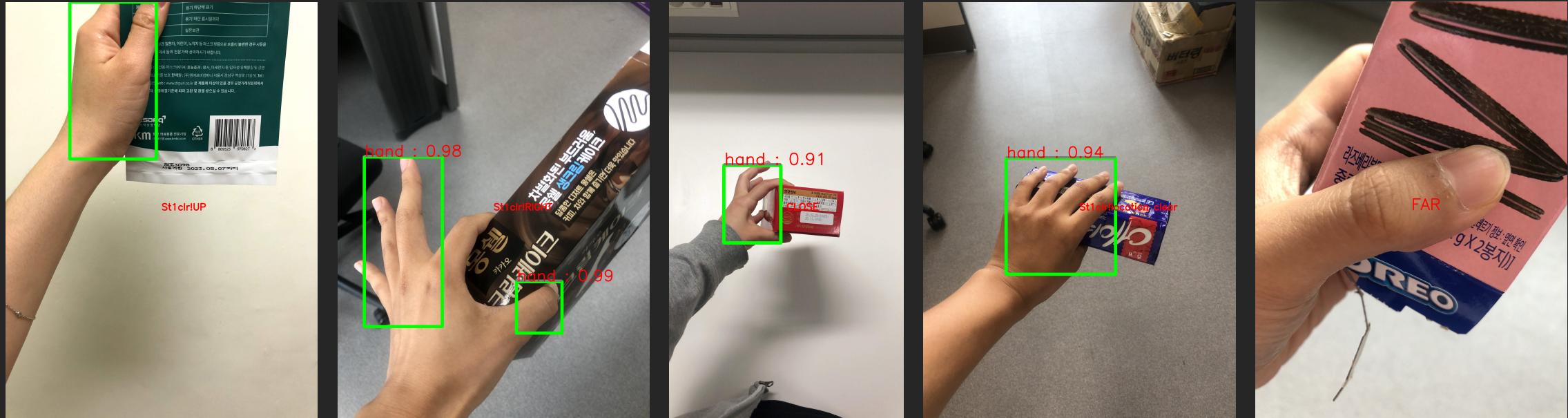
- Based on Evaluation,  
Task Completion Time and Accuracy of Each Task

- Can the tactile/auditory module provide accurate feedback to the user?  
→ Hand Feedback Accuracy / Performance

- Does this app help user to find the nutrition facts label and recognize them correctly?  
→ Nutrition Facts Extraction Accuracy / Performance

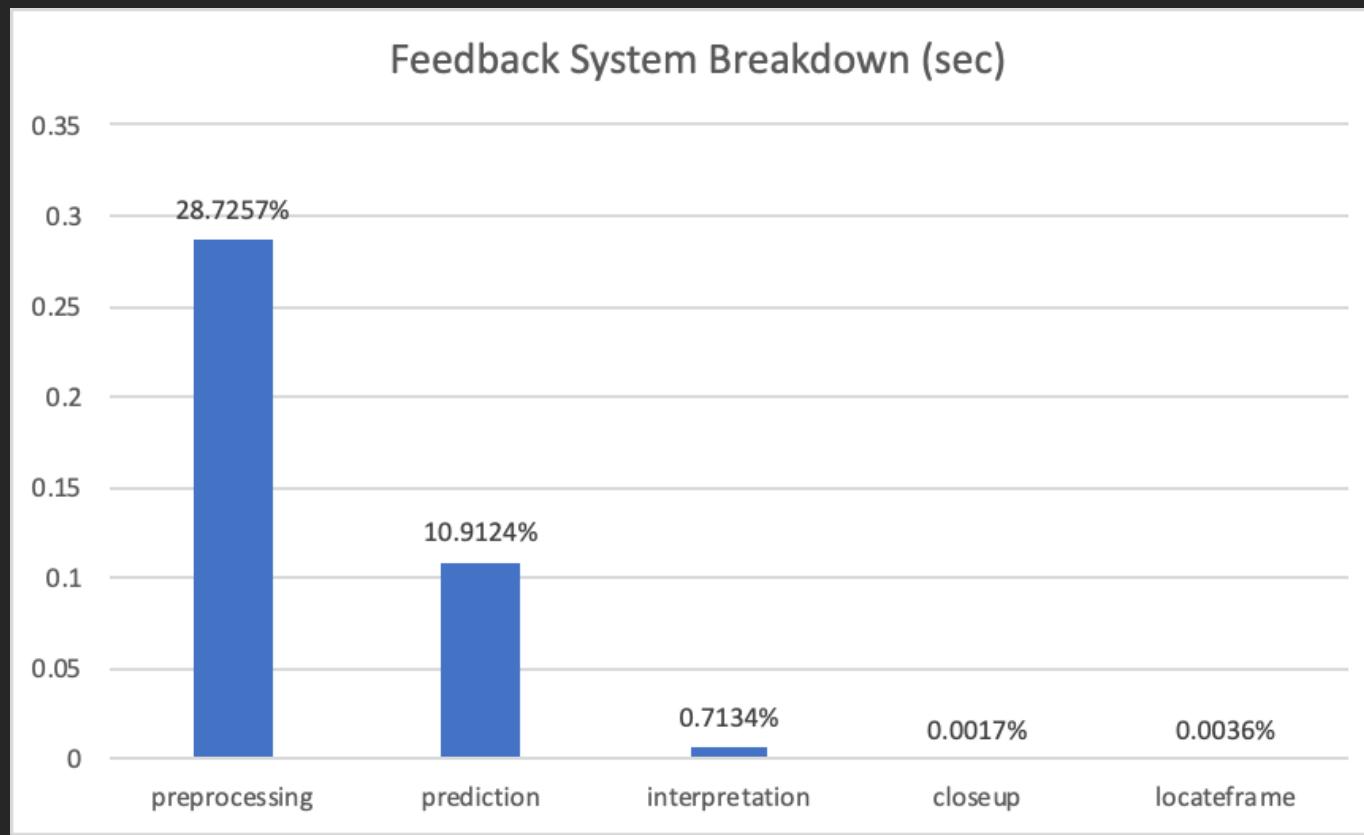
- Does this app provide the proper nutrition information for customized user preference?

# Feedback Algorithm Accuracy



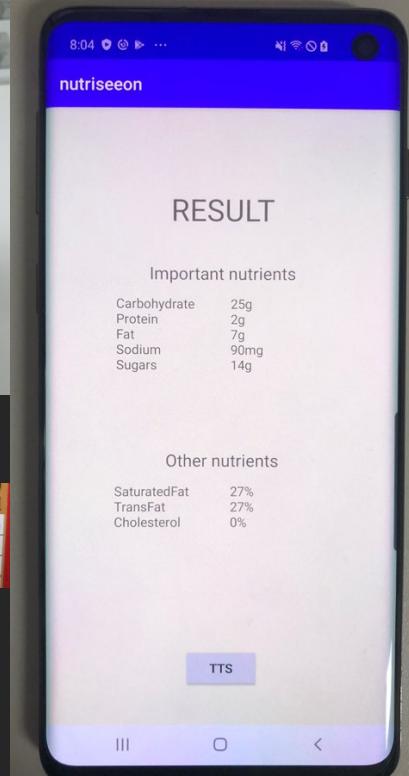
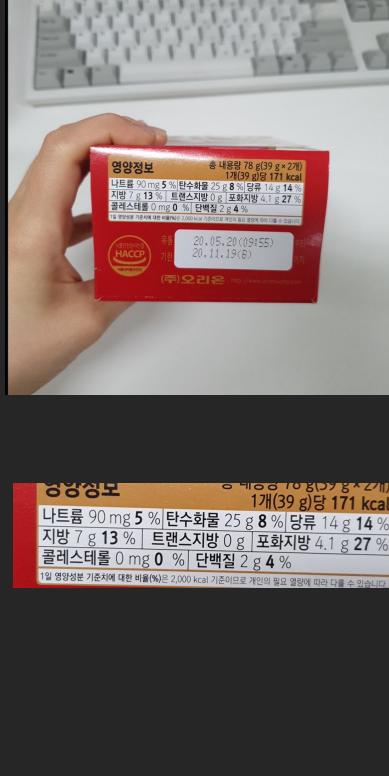
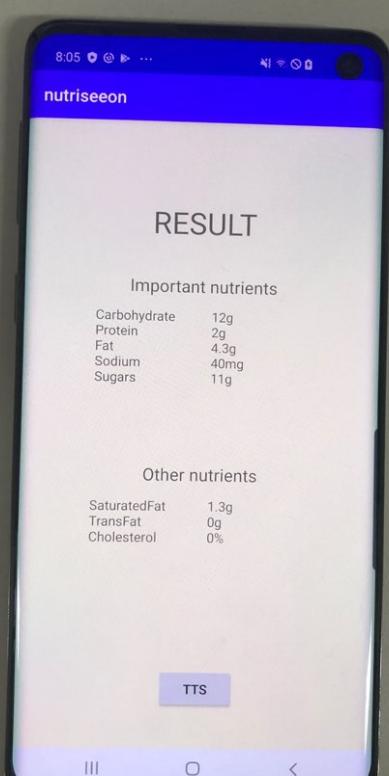
Our feedback module detects hands and provides the appropriate response for the location of the recognized hand.

# Feedback Module Performance

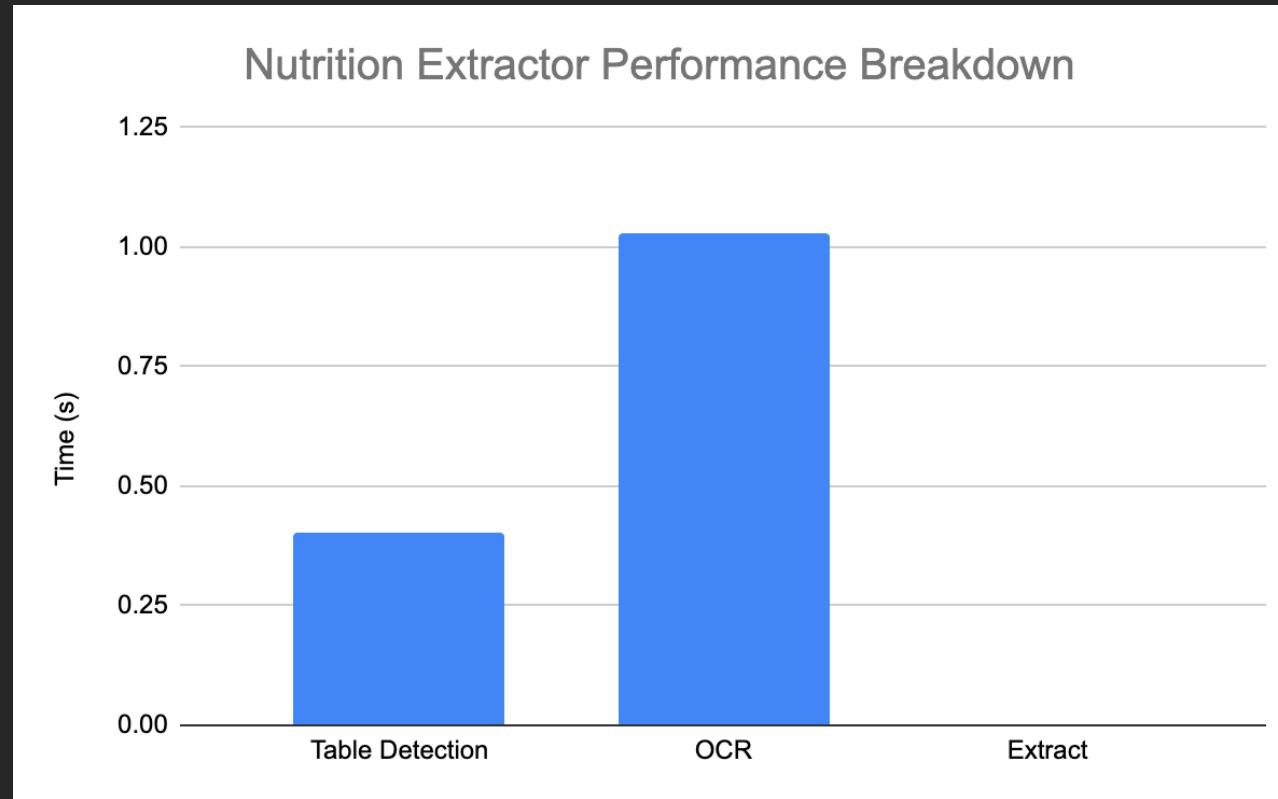


Total 0.403567502121466 s

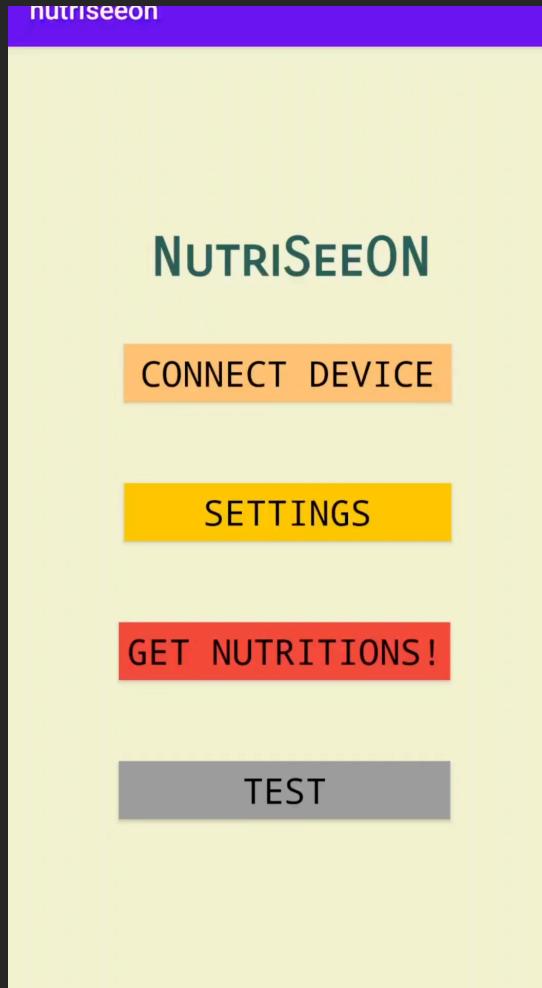
# Nutrition Extractor Accuracy



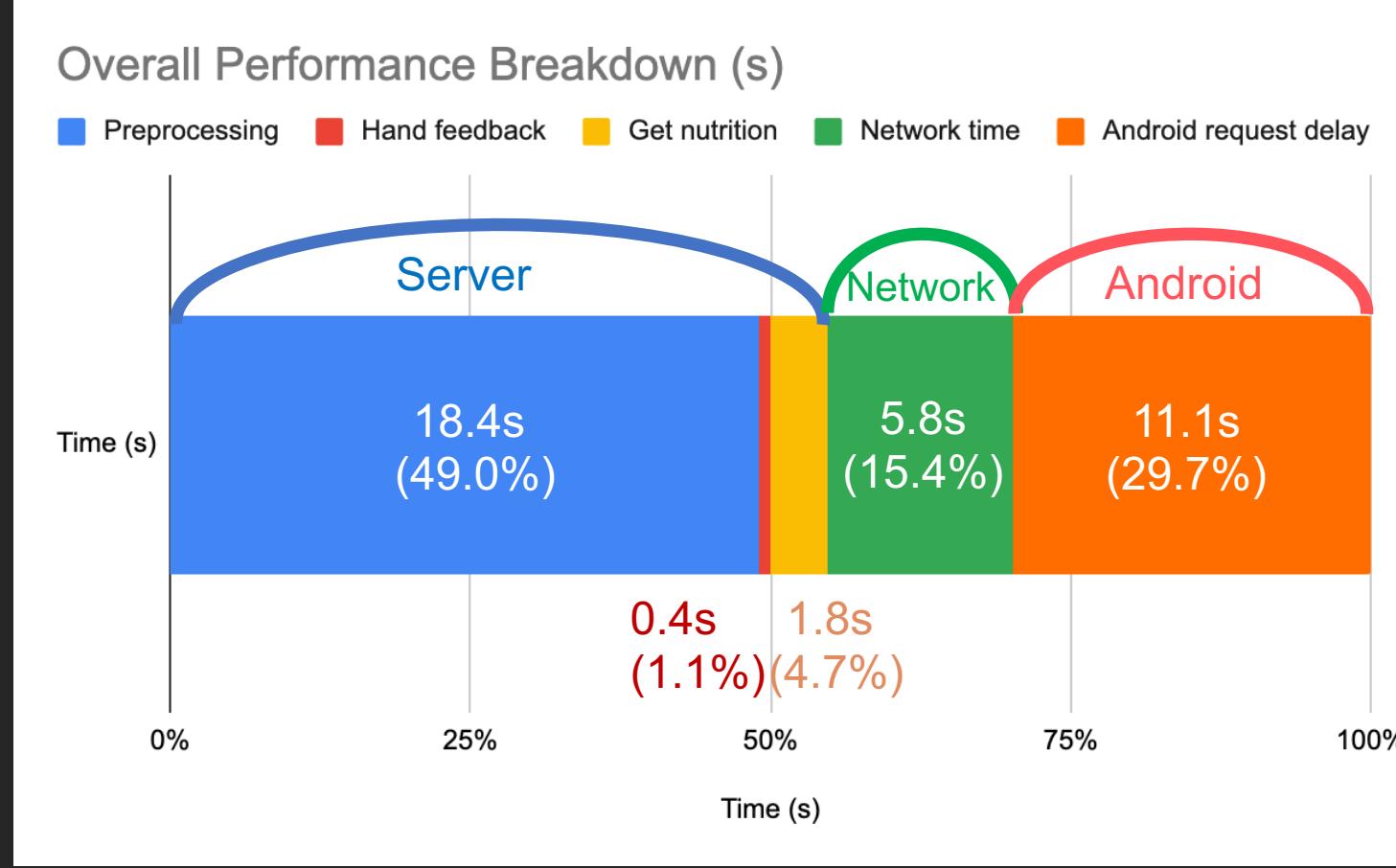
# Nutrition Extractor Performance



# Final Deliverable



# End-to-end Performance



# Future works

Our system support food products that is

- box-shaped, with a low reflective surface, on clean background
- and containing the formalized nutrition facts (either in English / Korean)
- Mobile device on right hand

Supported



We can extend our project to support more products

Unsupported



Reflective,  
not box-shaped



Nutrition facts table  
not formalized

# Project Management

- Eunsu
- Juhee
- Seokhyeon

# Timeline & Role



## Ideation

- Prior study
- Find public datasets / trained models

## Project Set-up

- Development environment setting (Client)
- Development environment setting (Server)
- Hand detection dataset construction

## Core Function Implementation

- Hand detection
  - Hand location feedback
  - Tactile feedback hardware impl.
  - Hand detection post-processing
- Nutrition Facts Table Recognition
  - OCR
  - OCR post-processing

## Application Development

- Client Implementation (Mobile application)
- Server Implementation (Network, Data Processing)

## Test and Evaluation

- Evaluation
- Test

# Thank you

**Nutrition Facts**

About 14 servings per container.

Serving size  
2 Tbsp (32g)Calories  
**200**

% Daily Value

Total Fat 16g  
21%Saturated Fat 3g  
15%

Trans Fat 0g

Cholesterol 0mg

4%

Sodium 90mg

0%

Total Carbohydrate 10g

4%

Dietary Fiber 2g

7%

Total Sugars 8g

Includes 6g Added Sugars 12%

Protein 5g  
3%

Vitamin D 0mcg 0% • Calcium 50mg 4%

from 0.7mg 4% • Potassium 90mg 2%

The % Daily Value (DV) tells you how much a nutrient in a serving of food contributes to a daily diet. 2,000 calories a day is used for general nutrition advice.

INGREDIENTS: Dry Roasted Almonds, Organic Powdered Sugar, Organic Cane Sugar, Organic Cocoa Butter, Palm Oil, Organic Vanilla Extract, Sea Salt.  
CONTAINS: Almonds. Produced on equipment that also processes other tree nuts.  
CONTAINS: Wheat, Soy, Milk, Egg, Peanuts, Tree Nuts, Sesame, Fish, Shellfish, Mustard, Lupine, Sesame, Sulfites, Yeast, Caffeine, Caramel Color, Artificial Flavors, Artificial Colors, and/or Preservatives.