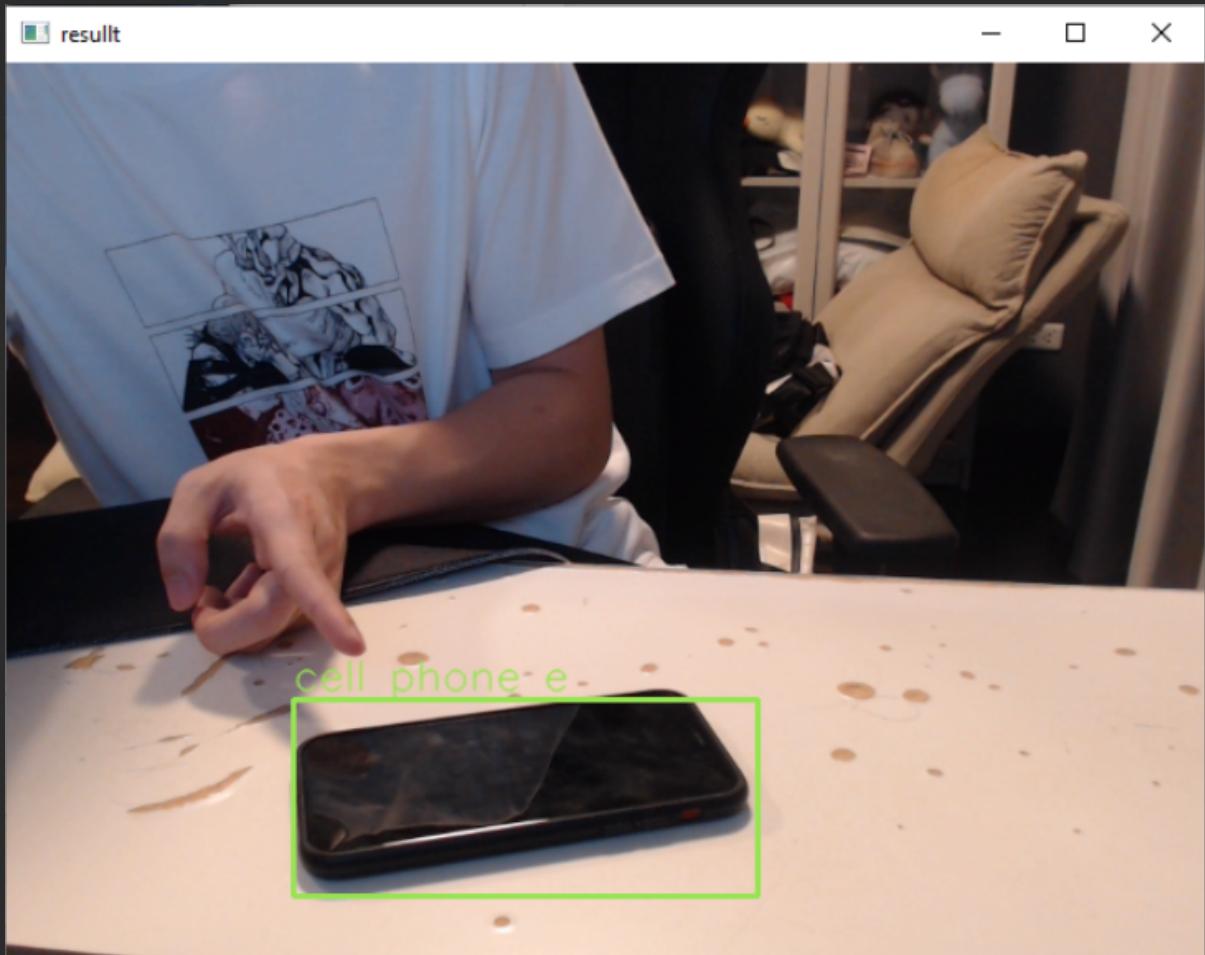


# Household object detection from finger tip



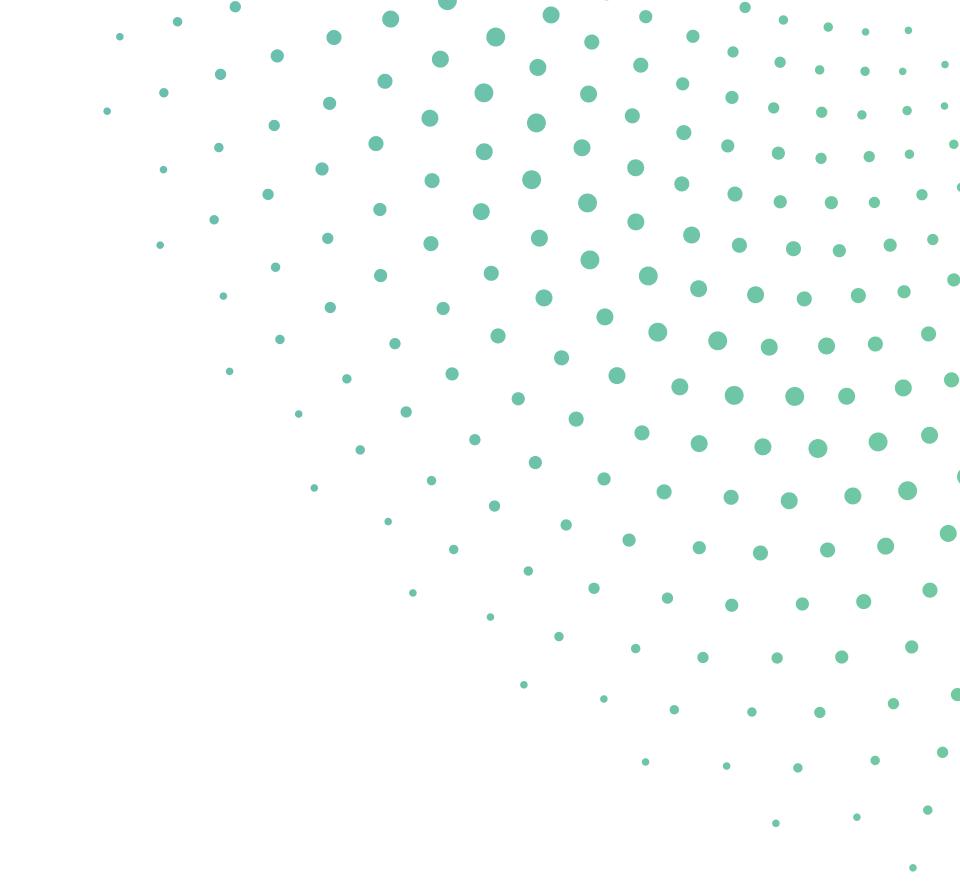
Group: Wall-E 2  
6330301321 Panithi Kamwangyang  
6330305921 Pras Pitasawad  
6330563421 Akira Sitdhikariyawat

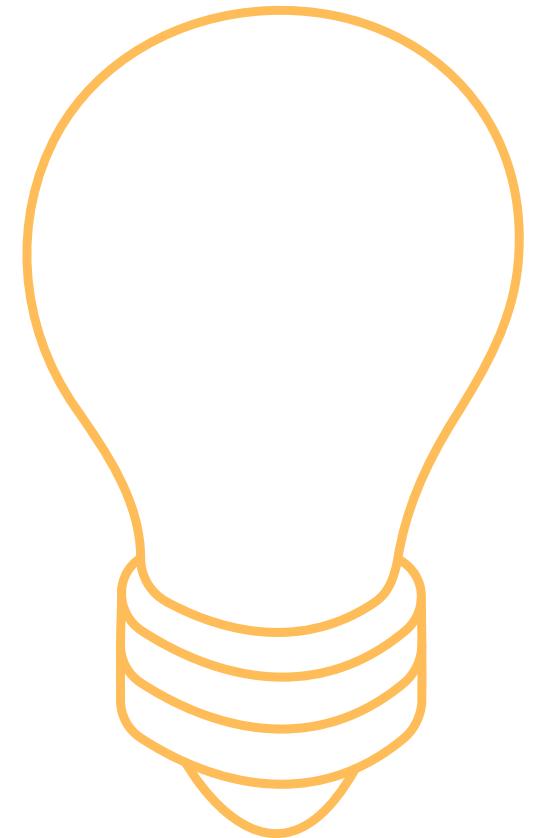
>>>

# Agenda



- Project scopes
- Challenges
- Project design
- 1st Approach:  
**Stereovision**
- 2nd Approach:  
**2 Camera setup**
- Demonstration





# Project scopes

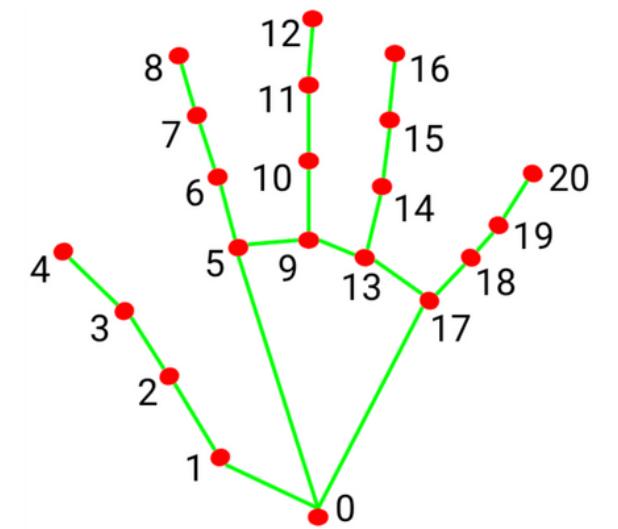
- Basic household object detection from right hand's index fingertip
- 2 camera in fixed place

## Challenges

- Stereovision is not simple
- Responsive output
- Robustness

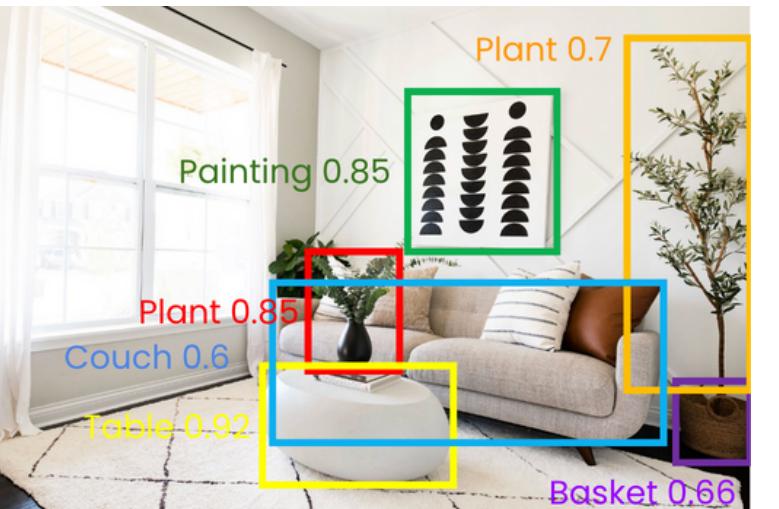


# Project design



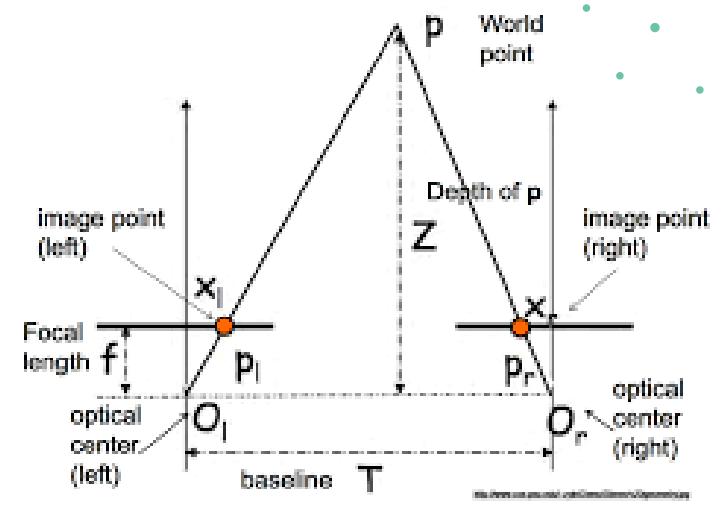
## Hand tracking

- Using MediaPipe's model to get the exoskeleton coordinate XY of the hand in the picture.



## Object detection

- Using MediaPipe's pre-trained object detector SSDMobileNet
- Basic household object ex.water bottle, laptop etc.



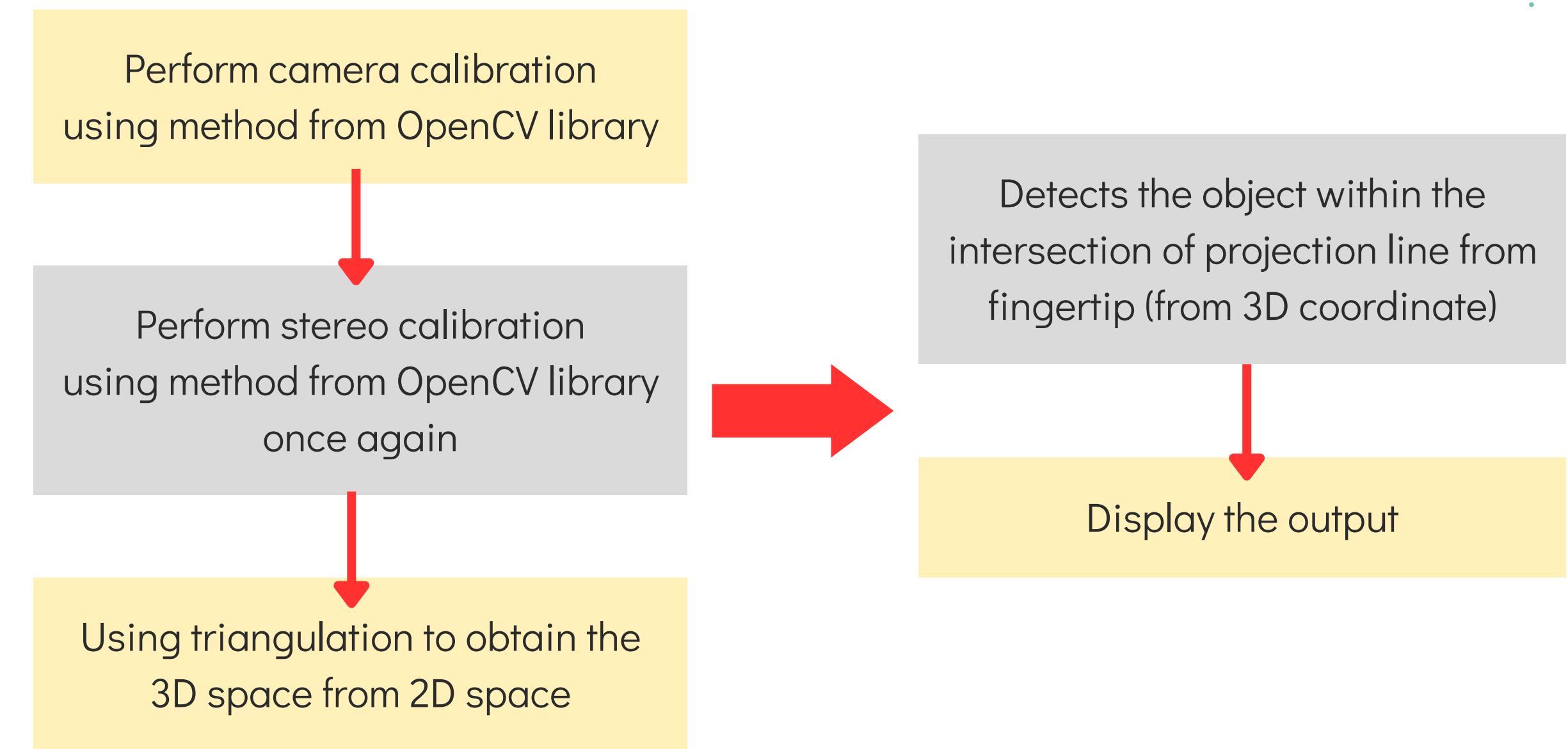
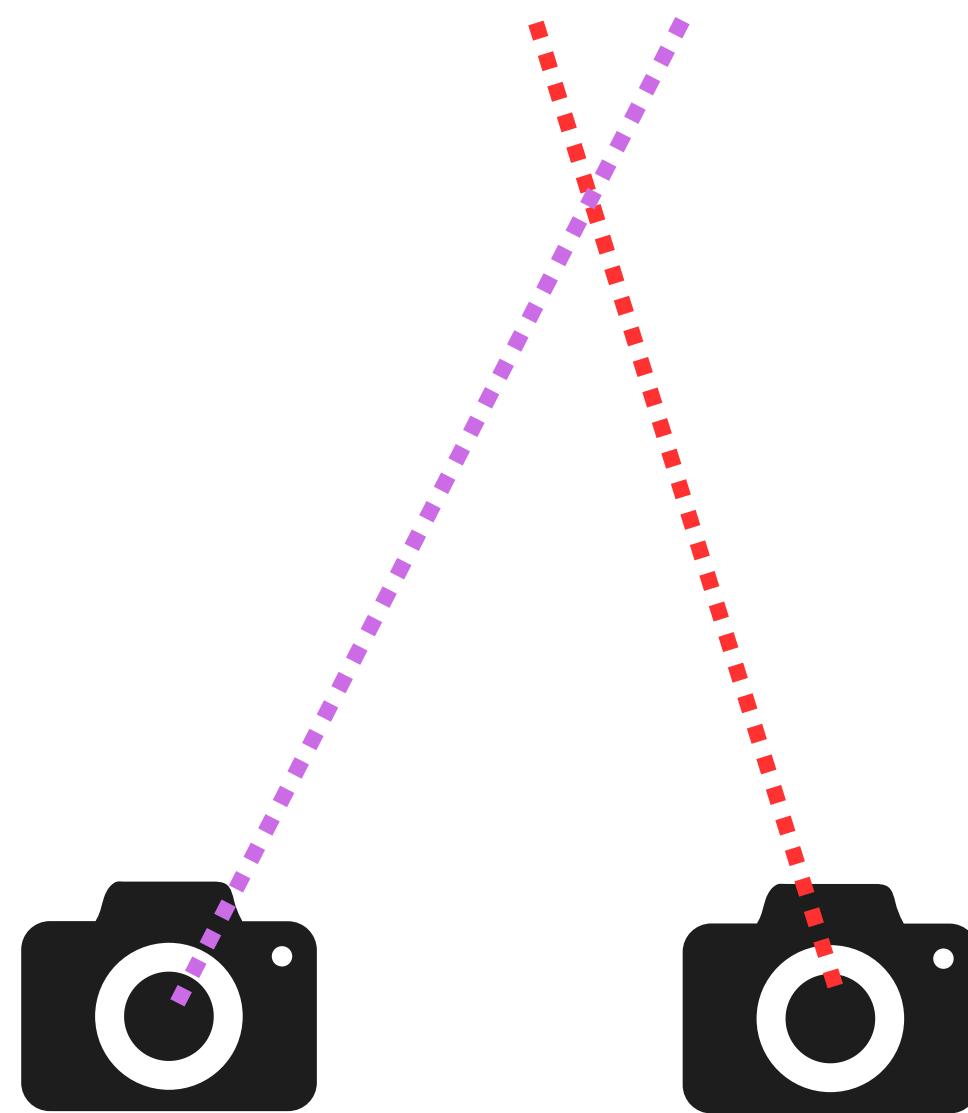
## Stereovision

- Calculate the projection line from the fingertip to identify which object is pointed.



# 1st approach: Stereovision

## Methodologies



# Stereovision

## What went wrong?

We were unable to correctly calculate the real world coordinate XYZ, which make it impossible to find the projection of the fingertip. This is crucial because without the coordinate XYZ, we can't identify which object is pointing to.

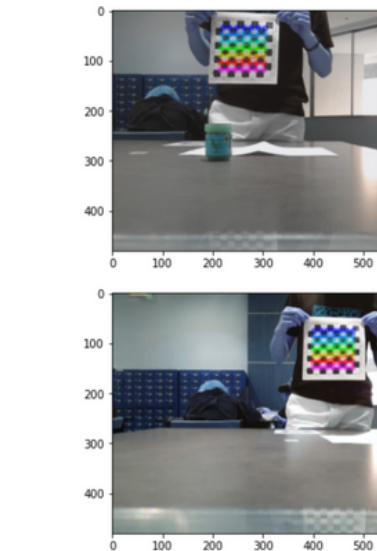
## Why?

### CAMERA SPECIFICATION

- Different lens, sensor depth, focal length etc. which makes the calculation not precise.
- Some information is unknown, ex. focal length - need to use estimate value

### POOR CAMERA CALIBRATION

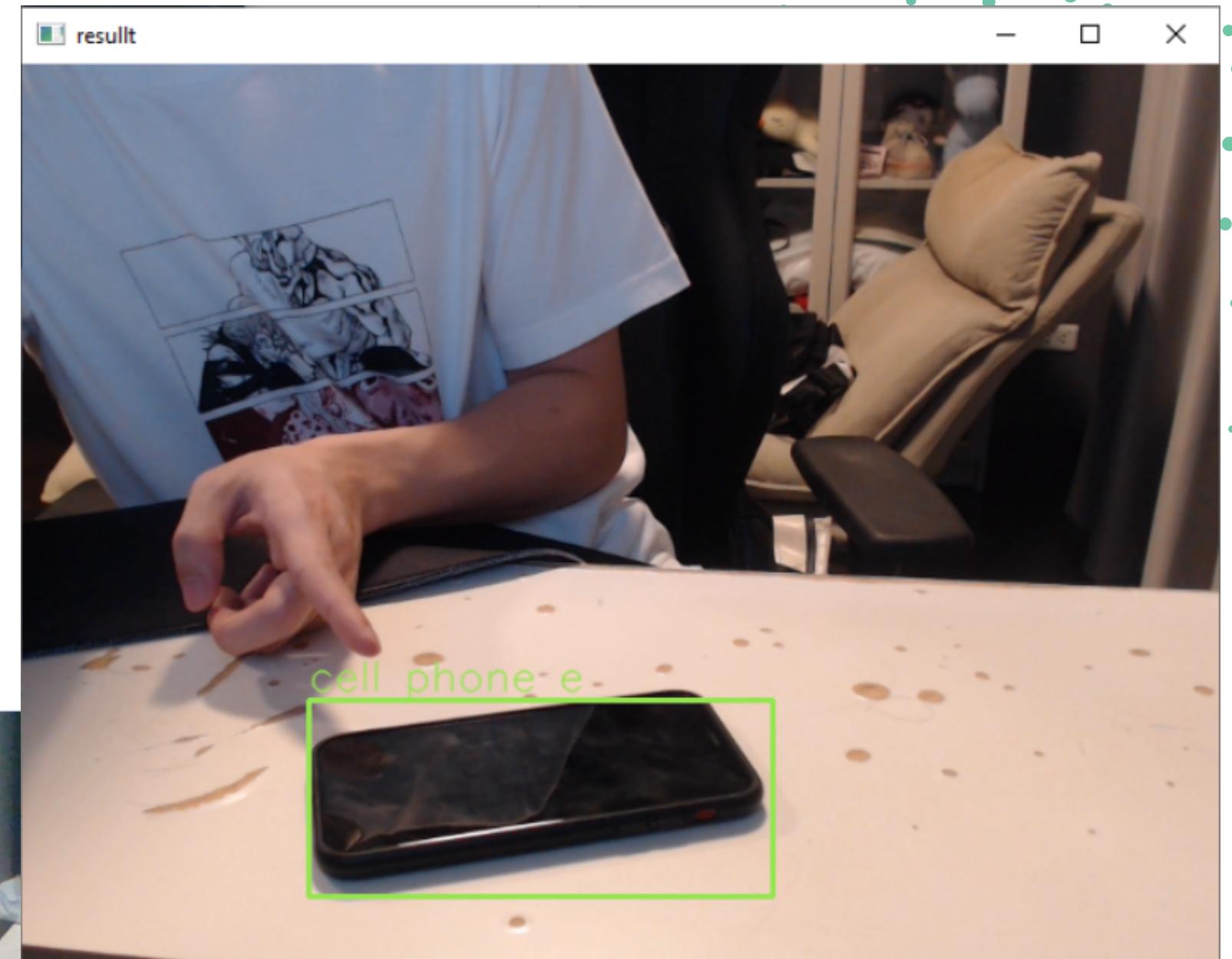
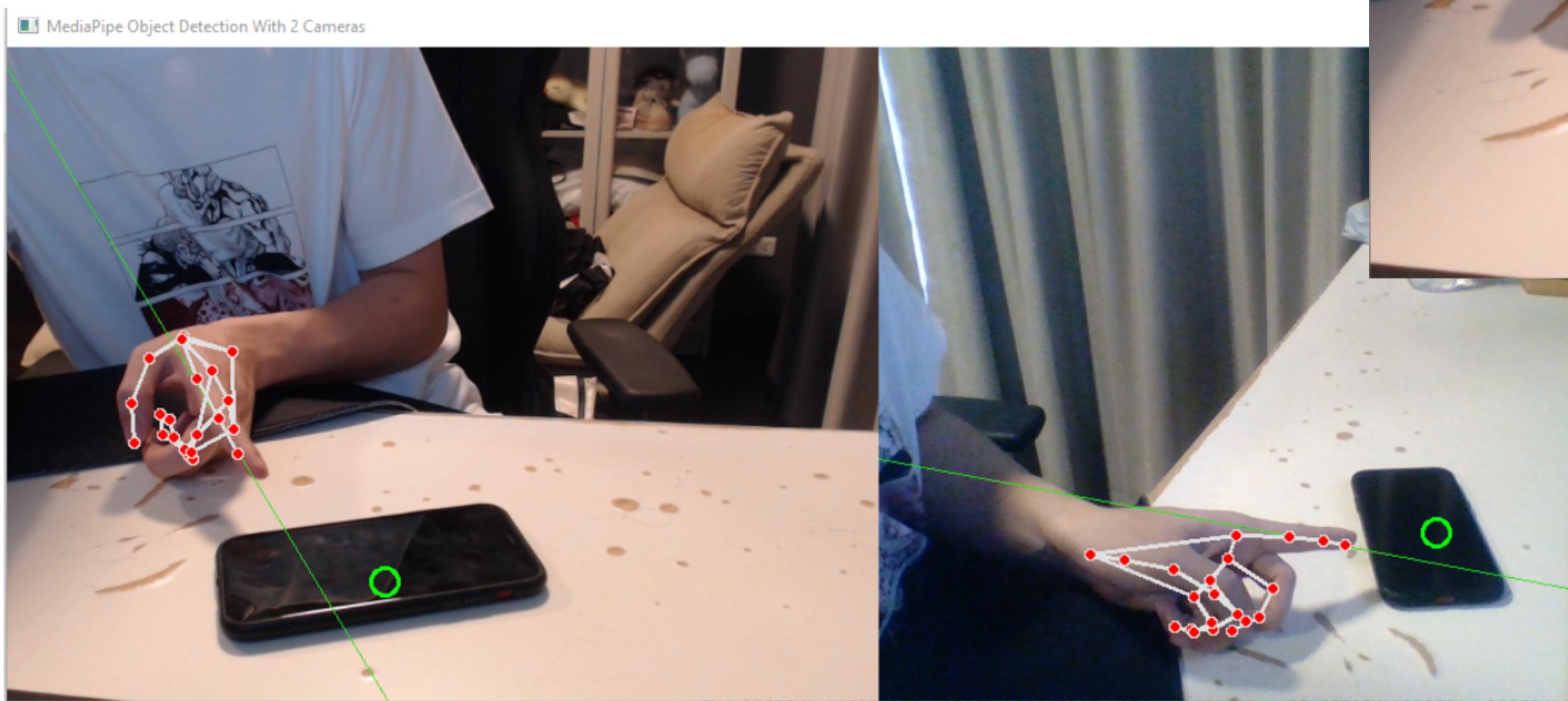
- Camera calibration plays key role in determining the camera matrix, without appropriately calibrating, the calculation will be error.



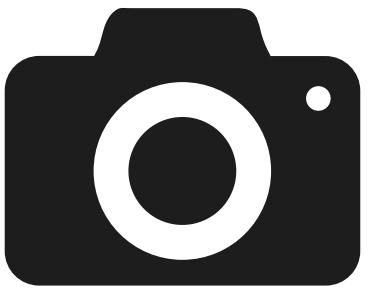
```
Left Camera Matrix:  
[[2.09145194e+03 0.0000000e+00 2.47251669e+02]  
[0.0000000e+00 1.90222685e+03 3.20180951e+02]  
[0.0000000e+00 0.0000000e+00 1.0000000e+00]]  
Right Camera Matrix:  
[[312.99757712 0. 328.0011193 ]  
[ 0. 312.93293435 246.53818365]  
[ 0. 0. 1. ]]  
Left Distortion Coefficients:  
[[-1.33277678e+00 -2.99244512e+01 8.09185345e-02 -2.20212788e-01  
9.49948013e+03]]  
Right Distortion Coefficients:  
[[ 2.97446562e-01 -2.40119377e+00 5.89258086e-03 2.29292743e-02  
8.60922641e+00]]  
Rotation Matrix:  
[[ 0.86739545 0.02191537 -0.49713665]  
[-0.03004417 0.99951362 -0.00835879]  
[ 0.49671166 0.02218643 0.867632 ]]  
Translation Vector:  
[[ 69.04622666]  
[ 6.90678362]  
[-76.226055 ]]
```

`(358.5, 238.0, 0.15507911392405063), (407.5, 238.0, 0.15507911392405063), (404.0, 302.0, 0.15507911392405063), (359.0, 300.5, 0.15507911392405063)`

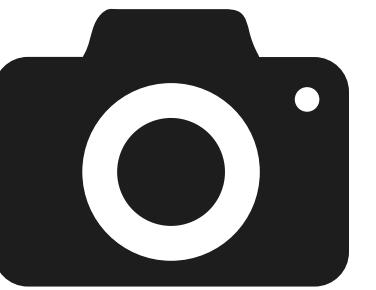
# 2nd approach: 2 Camera setup



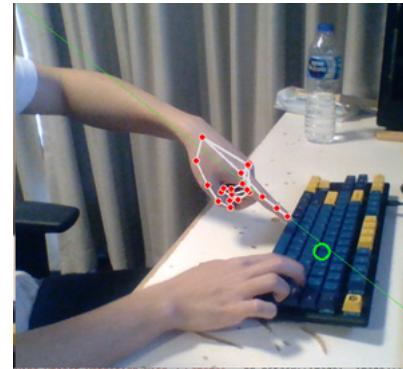
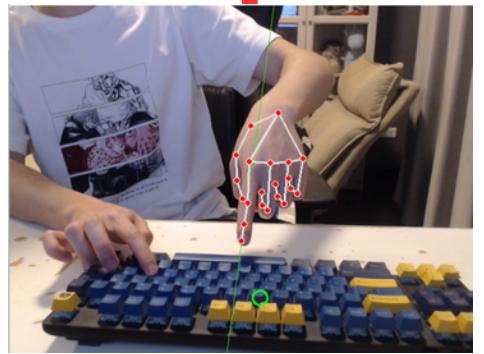
# 2 Camera setup



Main camera

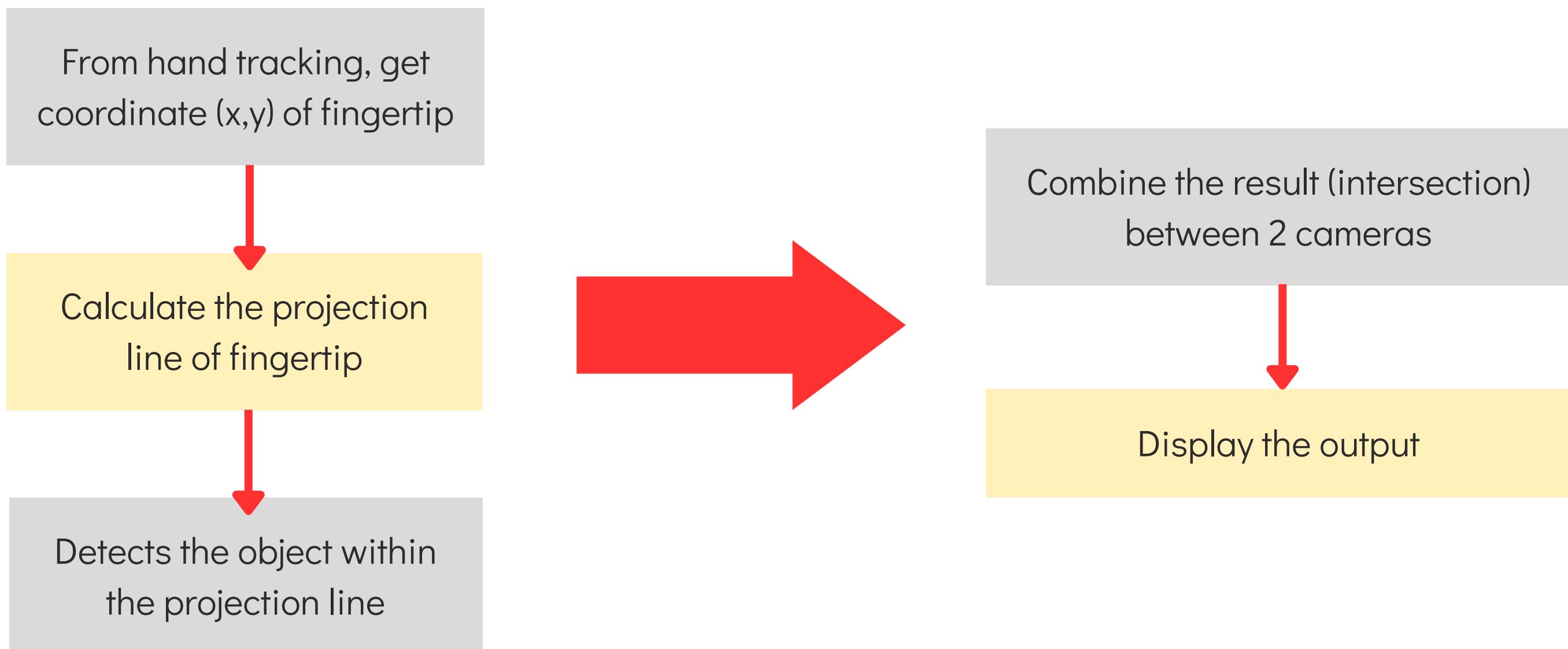


Second camera



# 2 Camera setup

## Methodologies



# Demonstration

