

Computer Science - Università degli Studi di Roma "La Sapienza"

# Pirate of the sea

Computer Vision

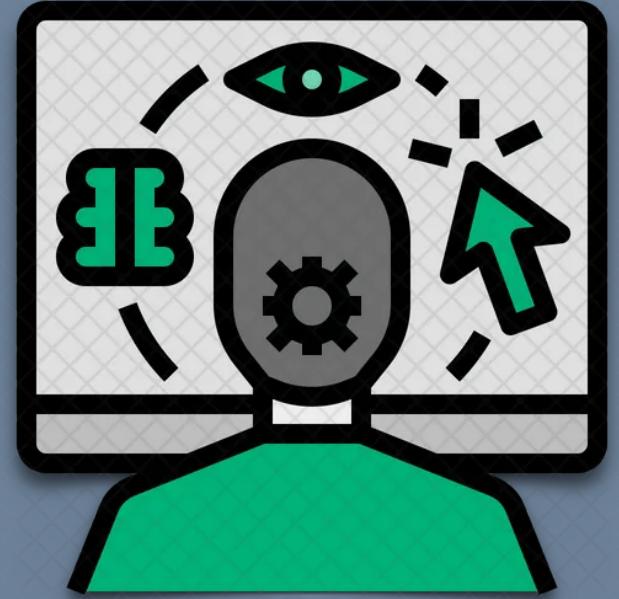
2023/2024

Alessio Buccarella - Riccardo Capobianco - Giorgio Scavalli Veccia

# The Paper

# Hand Gesture Recognition

Hand gesture recognition (HGR) is a fast-growing field of Computer Vision in several applications:



Human-computer  
interaction



Gaming



Virtual reality

# HGR Techniques

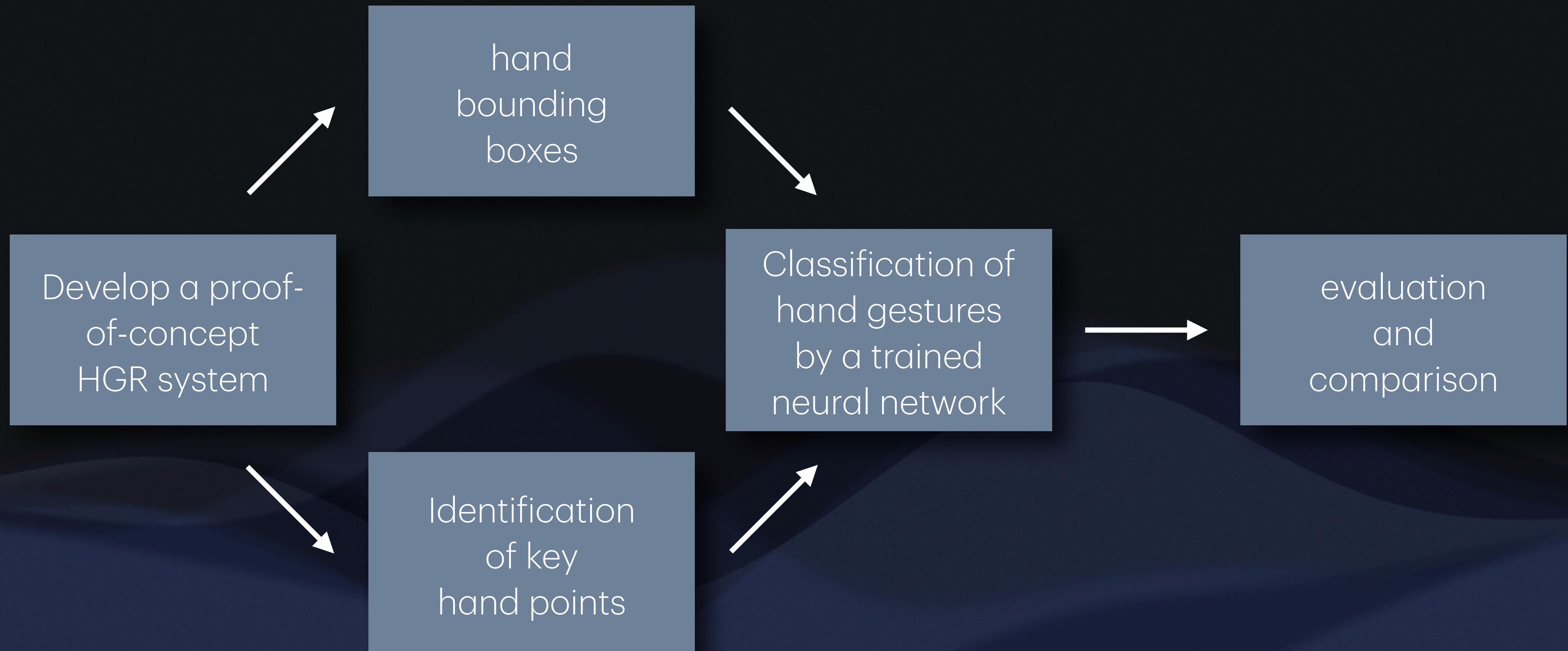
There are three main hand gesture recognition techniques:

- Template Matching
- Feature-based Approaches
- Deep Learning-based Approaches

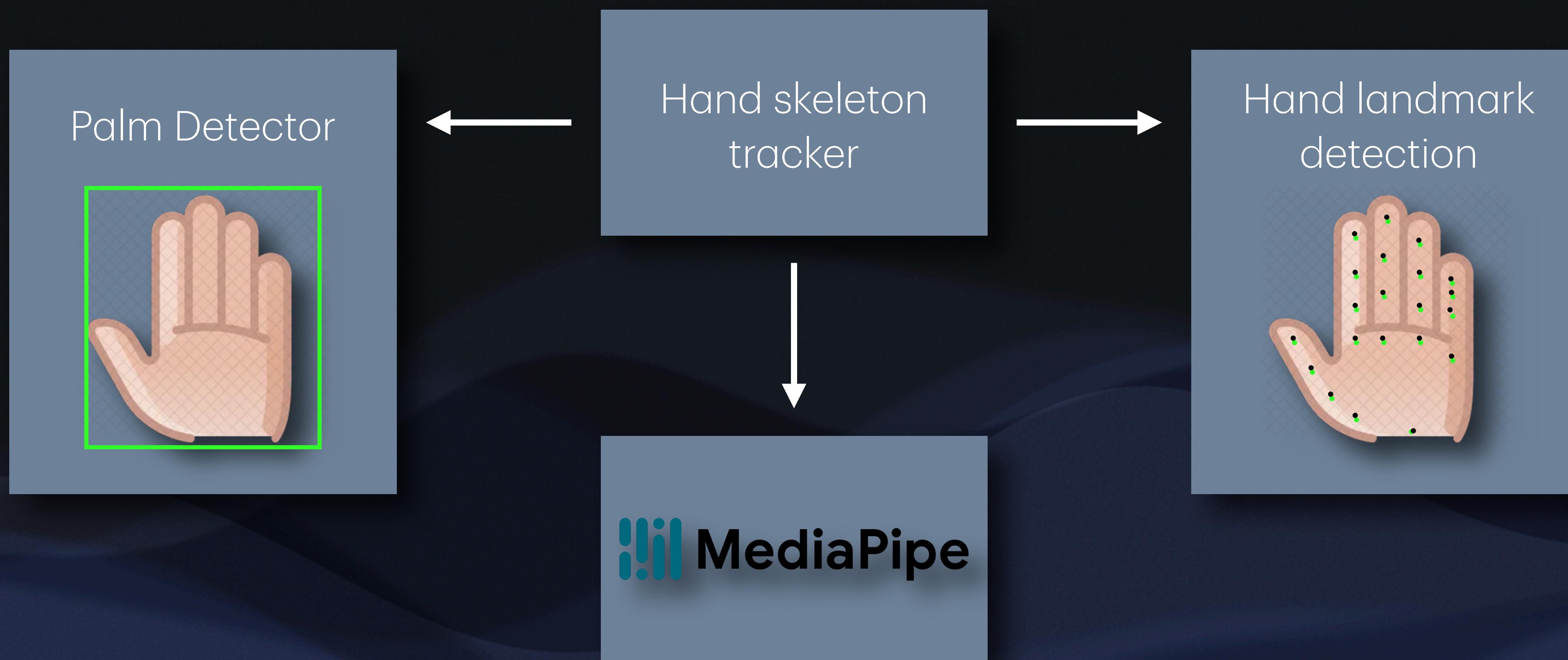
Hand gesture recognition technique used in the paper:

- Hand bounding boxes
- Identification of key hand points

# Paper mind map



# Methodology



# MediaPipe

- MediaPipe is an open-source development framework developed by Google, designed to simplify the development of **computer vision** applications.
- It provides **pre-trained models** to achieve accurate results in recognition and tracking activities of various elements in real time: hands (specifically palms), faces, human poses and other.

In particular:

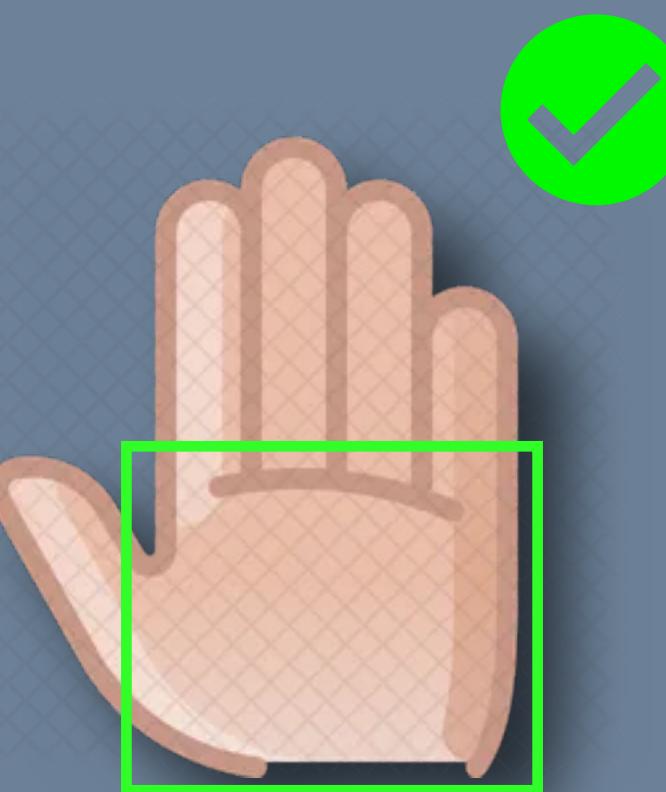
- Extensive **hand tracking** capability, with the ability to detect and track hands in various environments and orientations.
- **Extracts key points** (landmarks) on the hands, enabling detailed recognition of hand.

# Palm Detector

- Why Palm Detector?



Hand bounding boxes



Palm bounding boxes

# Hand landmark detection



# Gesture Classification Module

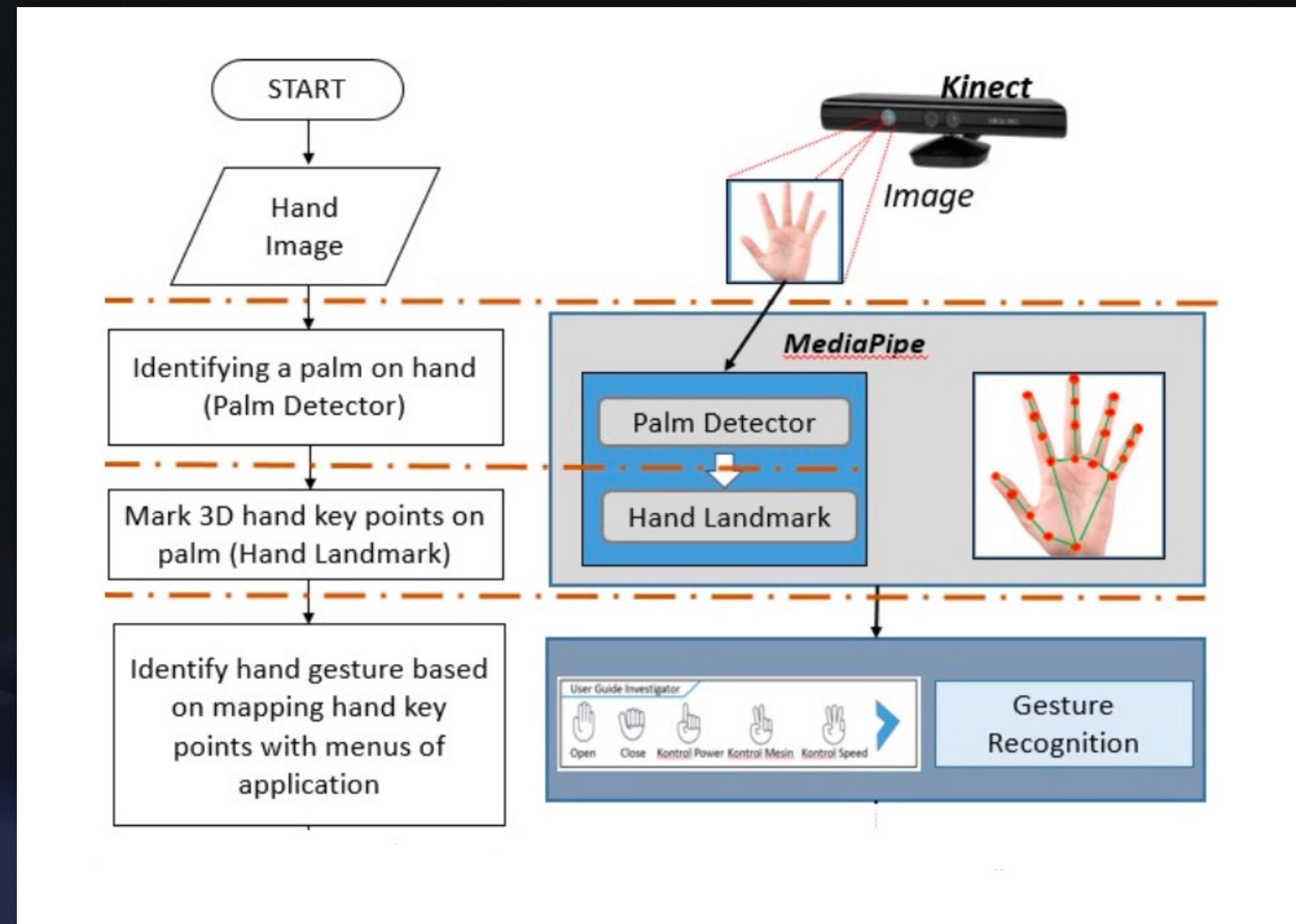
- The most essential information about the gestures might be present in the keypoints.
- So, how to classify the hand gesture according to key points?

## **Fully-connected (FC) neural network (NN):**

1. Each neuron in the network receives input from all neurons in the previous layer
2. Calculates a weighted sum of the inputs
3. Applies an activation function to the sum
4. Passes the output to neurons in the next layer

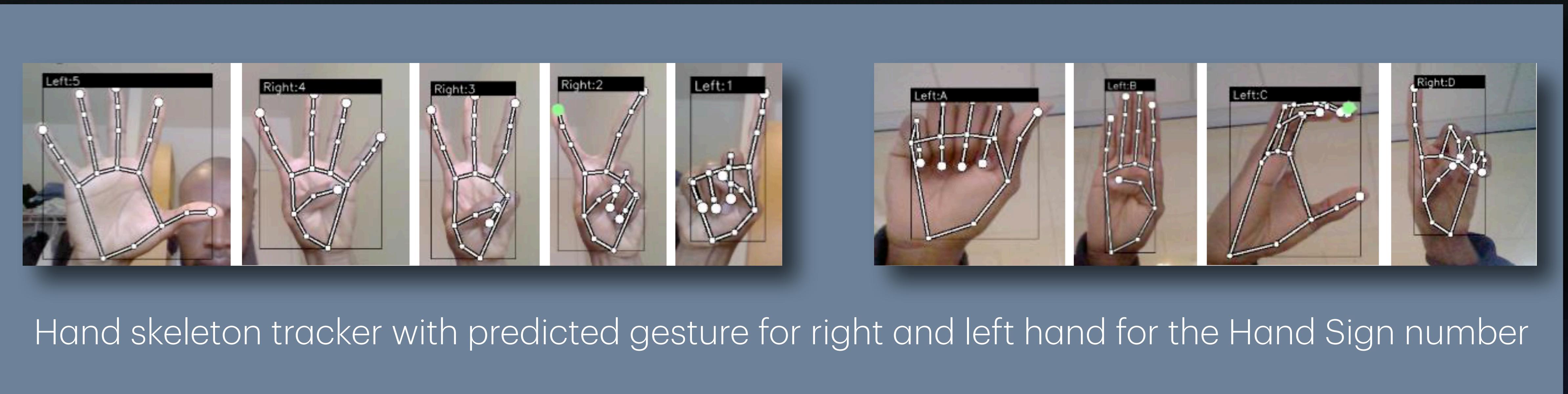
$$y_i = f(\sum_{j=1}^n w_{ij} x_j + b_i)$$

# Workflow of the Gesture Recognition



# Result

- 0.76 accuracy for the HSN (Hand Sign Number) gesture dataset
- 0.77 for the ASL (American Sign Language) gesture dataset



# The Project

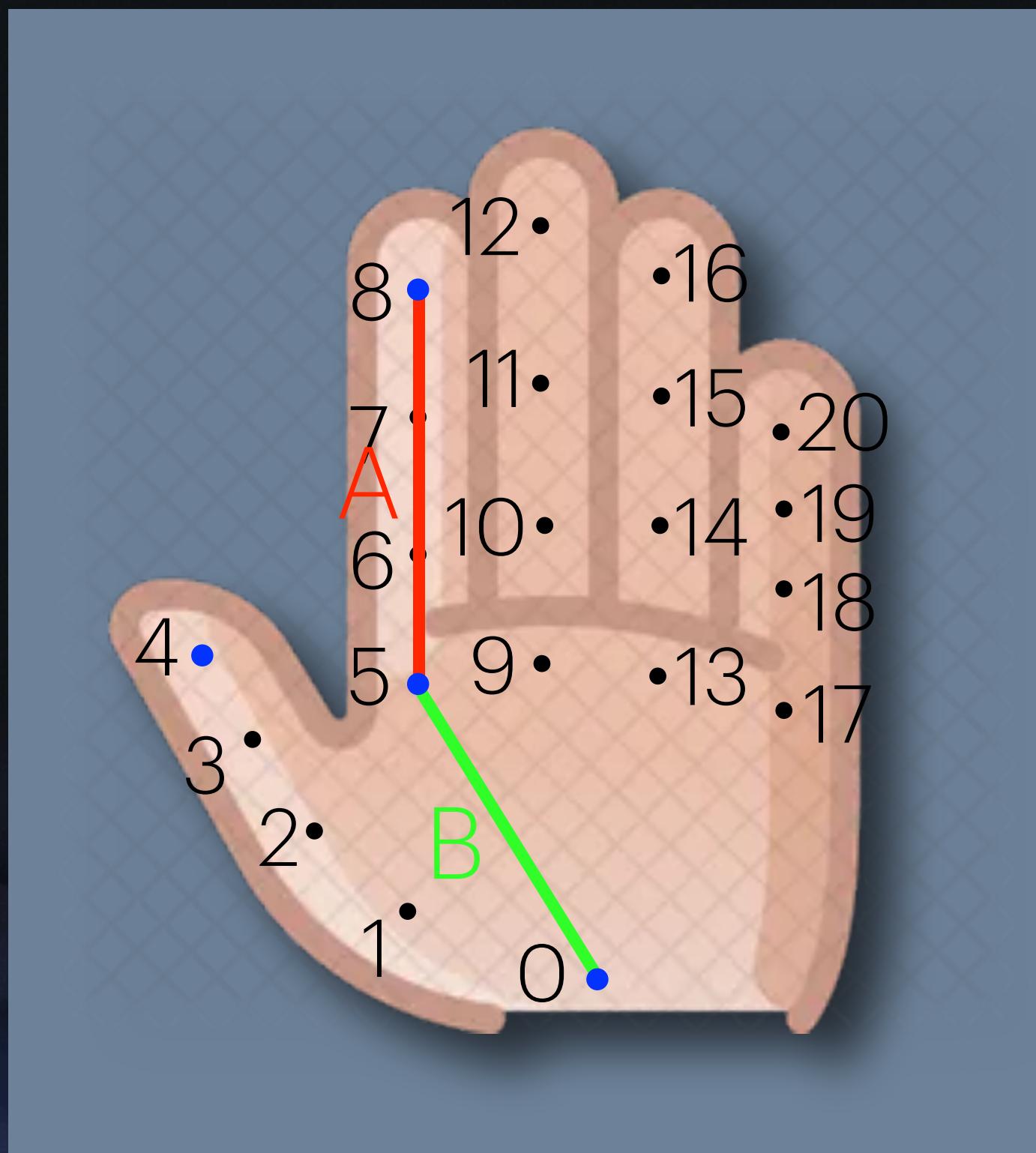
# Hand Pose Analysis (1)

- Using key points we defined disequations to track them and understand the pattern for opened/closed hands.
- We defined 2 different check in order to include more hand positions scenarios:

```
open=True
for i in range(1,5):
    if ((hand_tips[i].y-hand_bottom[i].y)> (hand_bottom[i].y -wrist.y - threshold_general) and
        (hand_tips[i].y-hand_bottom[i].y)< (hand_bottom[i].y -wrist.y + threshold_general)):
        open=True
    else:
        open=False

if (open and ((hand_tips[0].y )> (hand_bottom[1].y - threshold_thumb)) and
    (hand_bottom[0].y < wrist.y - threshold_thumb_2) and
    (hand_tips[0].y - hand_tips[1].y)> threshold_thumb_3 )
```

# Hand Pose Analysis (2)



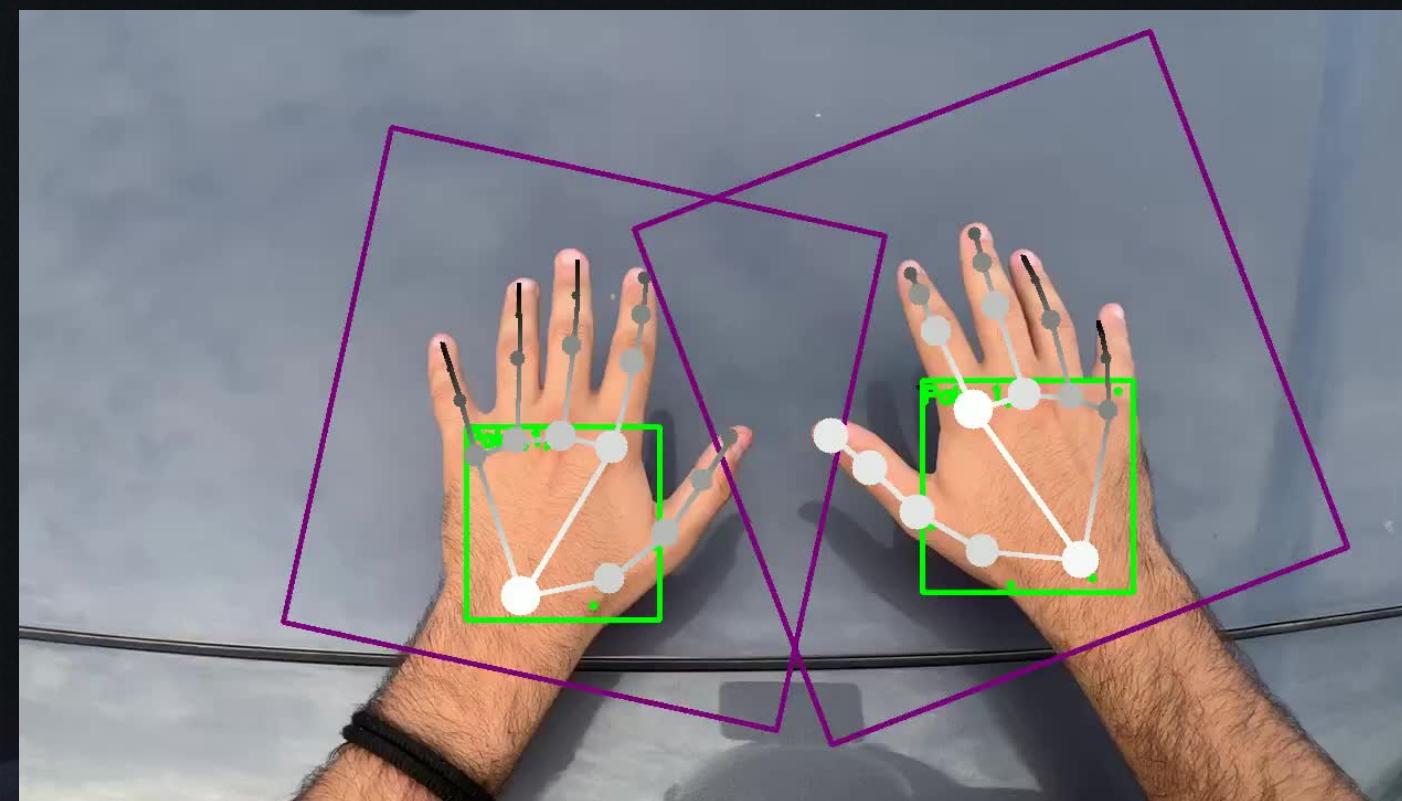
Finger X validation =  
$$\begin{cases} A.y > B.y - Th1 \\ A.y < B.y + Th1 \end{cases}$$

Hand first validation =  
Finger 1  
$$\begin{cases} \text{Finger 2} \\ \text{Finger 3} \\ \text{Finger 4} \end{cases}$$

Hand second validation =  
$$\begin{cases} \text{Thumb\_tip.y} < \text{Finger\_bottom.y} + Th2 \\ \text{Thumb\_tip.y} > \text{Finger\_bottom.y} - Th2 \\ (\text{Thumb\_tip.y} - \text{Finger1\_tip.y}) < Th3 \end{cases}$$

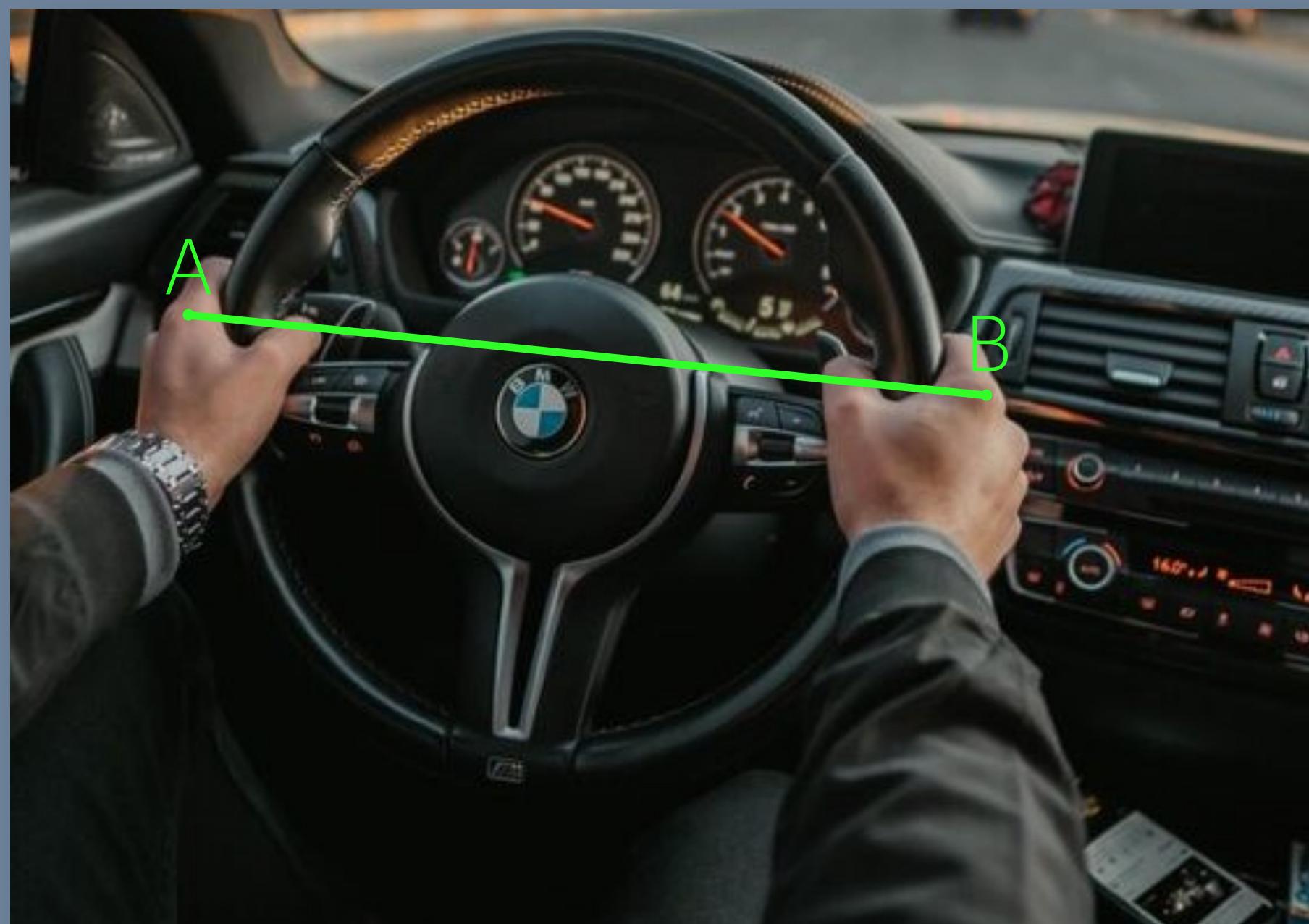
# Hand Pose Analysis (3)

- Logics to classify left and right hand:



```
if hand_tips[0].x < hand_bottom[1].x:  
    left_hand_status="Open"  
    left_hand_tips      = hand_tips  
    left_hand_bottoms  = hand_bottom  
    thumb2= (int(hand_bottom[1].x * frame.shape[1]),  
    int(hand_bottom[1].y * frame.shape[0]))  
    else:  
        right_hand_status="Open"  
        right_hand_tips      = hand_tips  
        right_hand_bottoms  = hand_bottom  
        thumb1= (int(hand_bottom[1].x * frame.shape[1]),  
        int(hand_bottom[1].y * frame.shape[0]))
```

# Hand rotation detection



- Compute hand rotation degrees using the arctangent function:

$$\arctan(yB - yA, xB - xA)$$

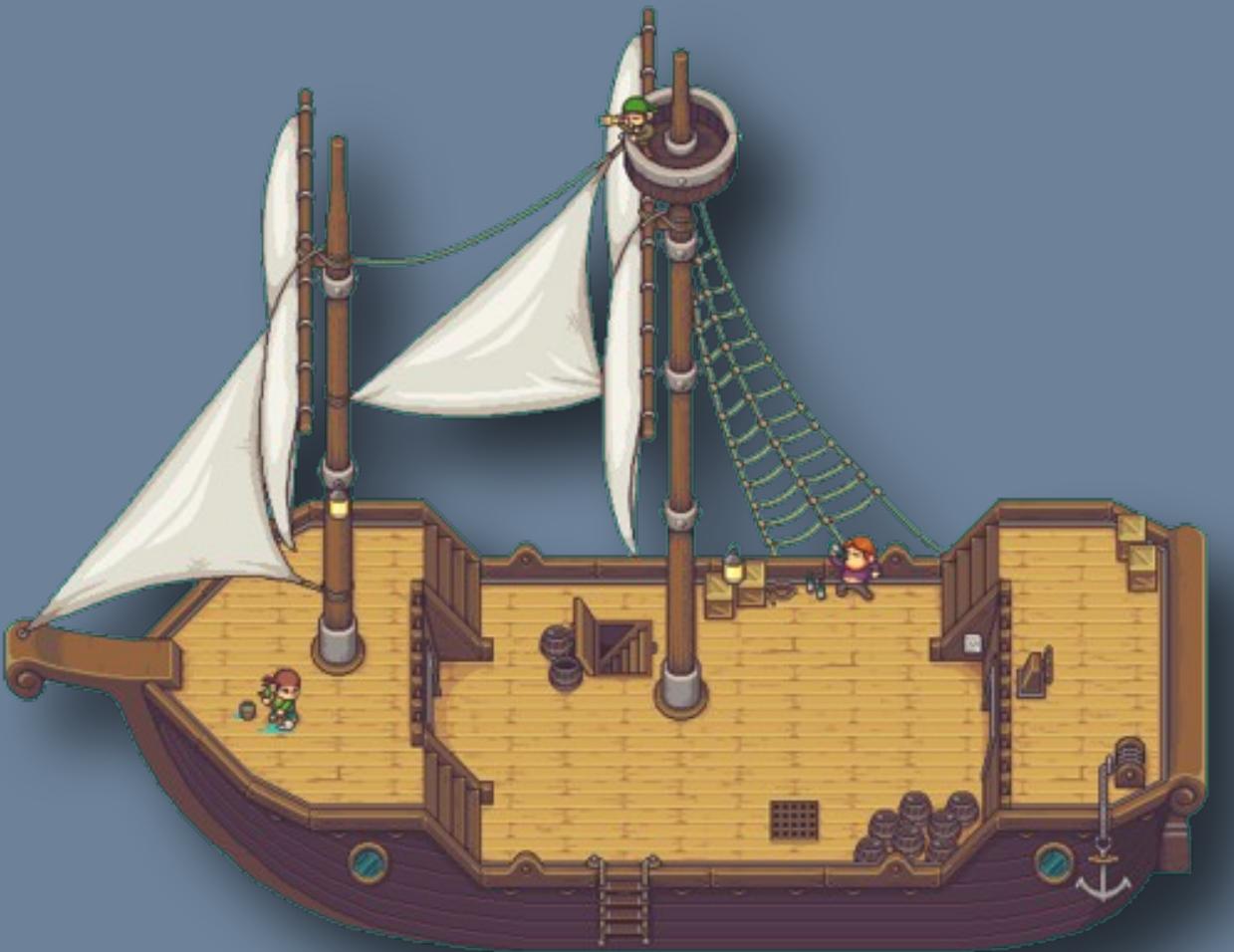
# Real-time camera

- Each frame captured by the camera is processed to improve the game experience.
- The game responds in real-time to the player's actions.



# The Game

- Main character:



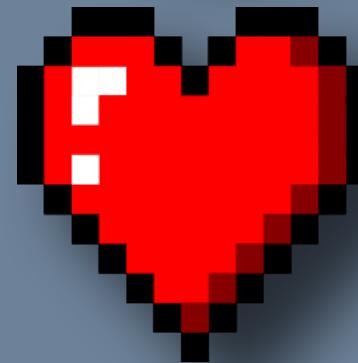
- Objects:



Avoid!



Catch!

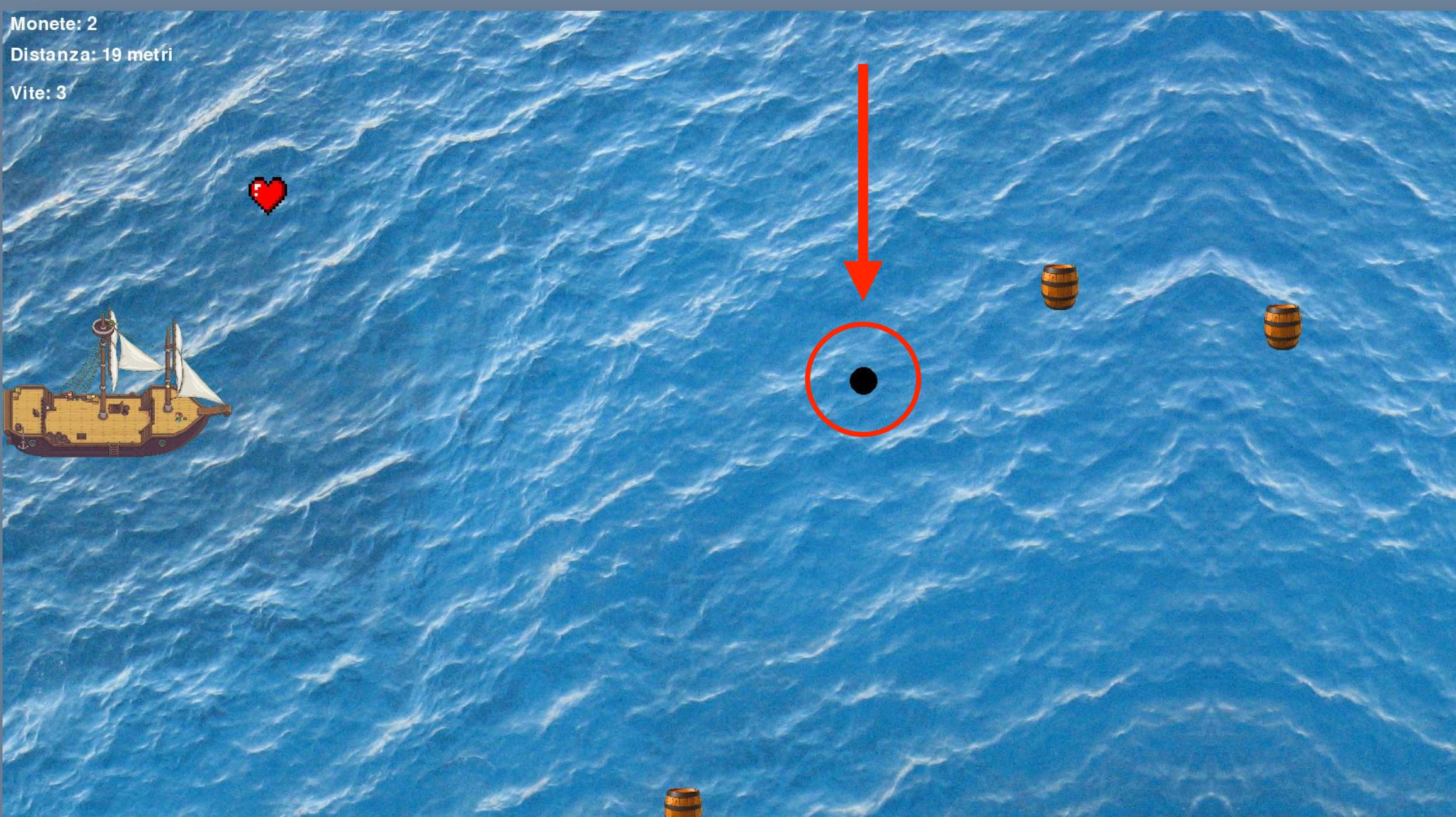


Catch!

# Open Hand



Shoot!



# Closed Hands



Drive!



# Game's goal

- Dominate the sea reaching the greatest possible distance
- Catch coins to become the richest pirate of the ocean



Thanks for your attention!