AI Coach for divers

Lekang Yuan and Xiaohang Yu

Lab of smart imaging, TBSI, Tsinghua University, China

TBSI

Abstract

- Apply the idea of smart coaching to a noval application: training of the neutral boudancy skill, a primary skill for scuba divers.

- Extract breath and motion states of the diver with underwater camera with CNN.

- Build a gym environment and train agents to determine the right action based on current states.

- The framework can be used to assess diver's skill, understand the dynamics, and provide real-time feedback.

Introduction
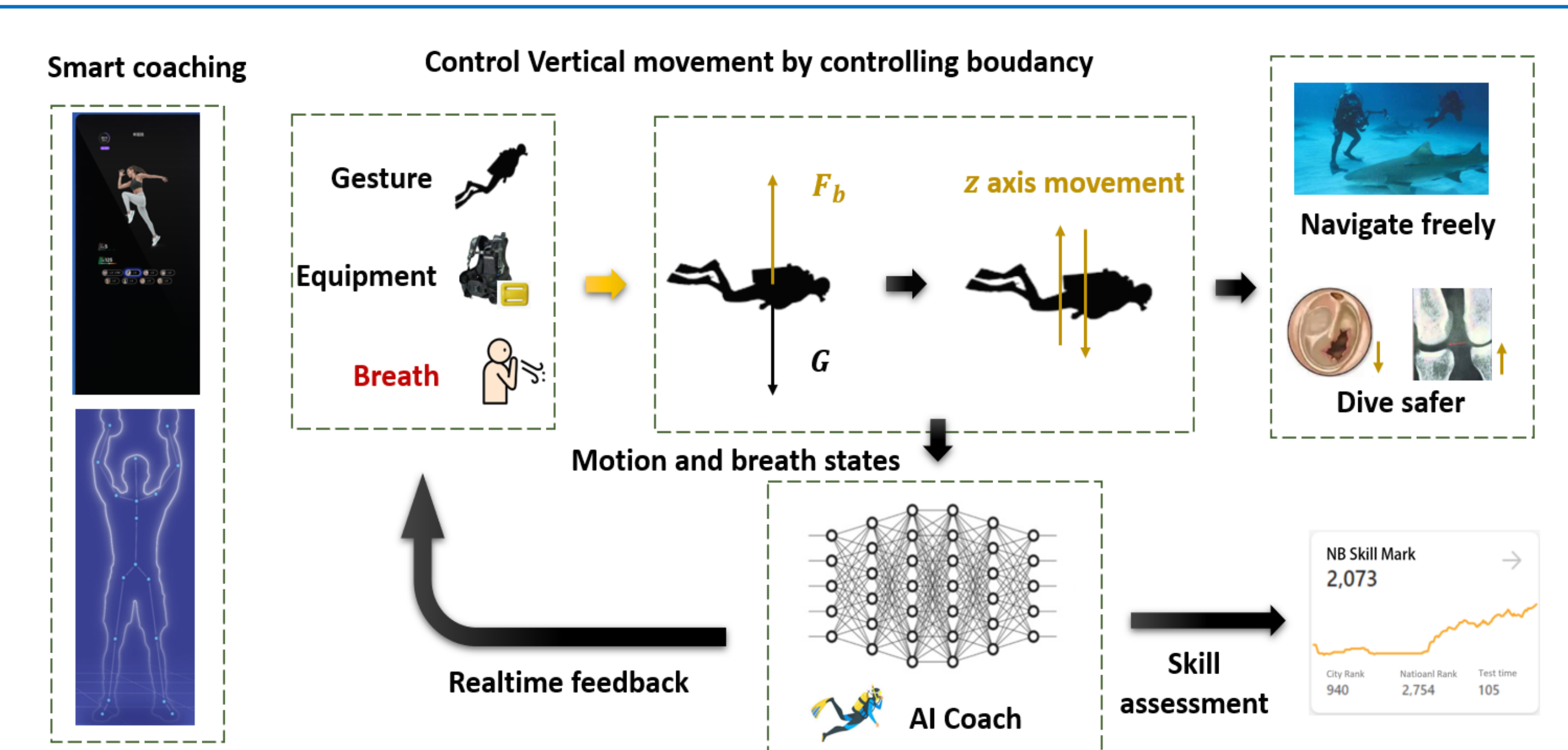
Smart coaching: Costless AI coach for accurate assessment and feedback[1].

Neutral boudancy skill: a primary skill for scuba divers

- Vertical movement is controlled by controlling boudancy while diving.
- Keeping neutral boudancy is essential for navigating freely underwater and diving safer.

**Machine learning problems for building an AI coach for diver:**

- **Peception:** How to extract motion and breath states from visual input?
- **Behavior optimization:** How to choose the right action based on current motion and breath states?

References

[1] FITURE, https://www.fiture.com/cn/, 2021
[2] What and how well you exercised? An efficient analysis framework for fitness actions[J]. 2021
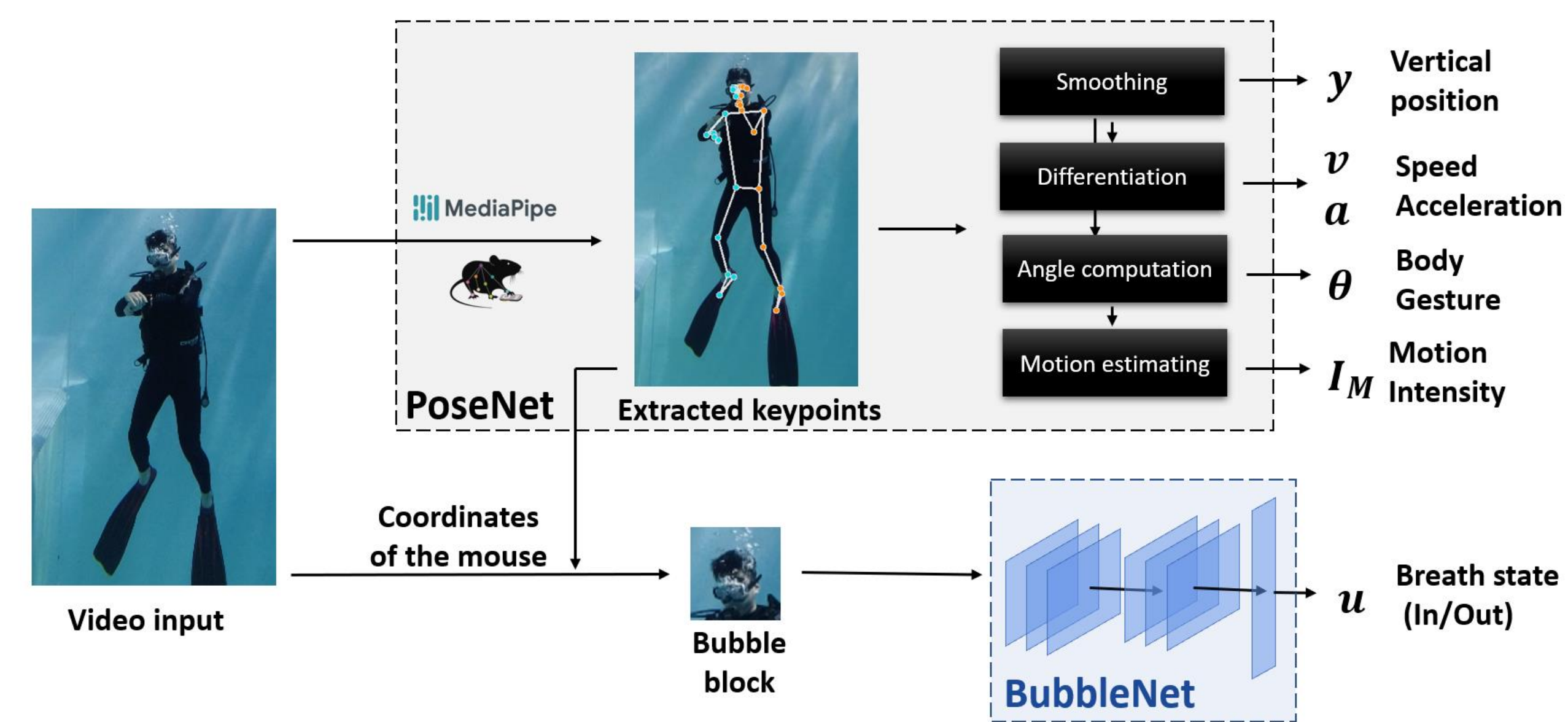[3] Mediapipe: A framework for building perception pipelines[J]. 2019.

Methods

Data collection
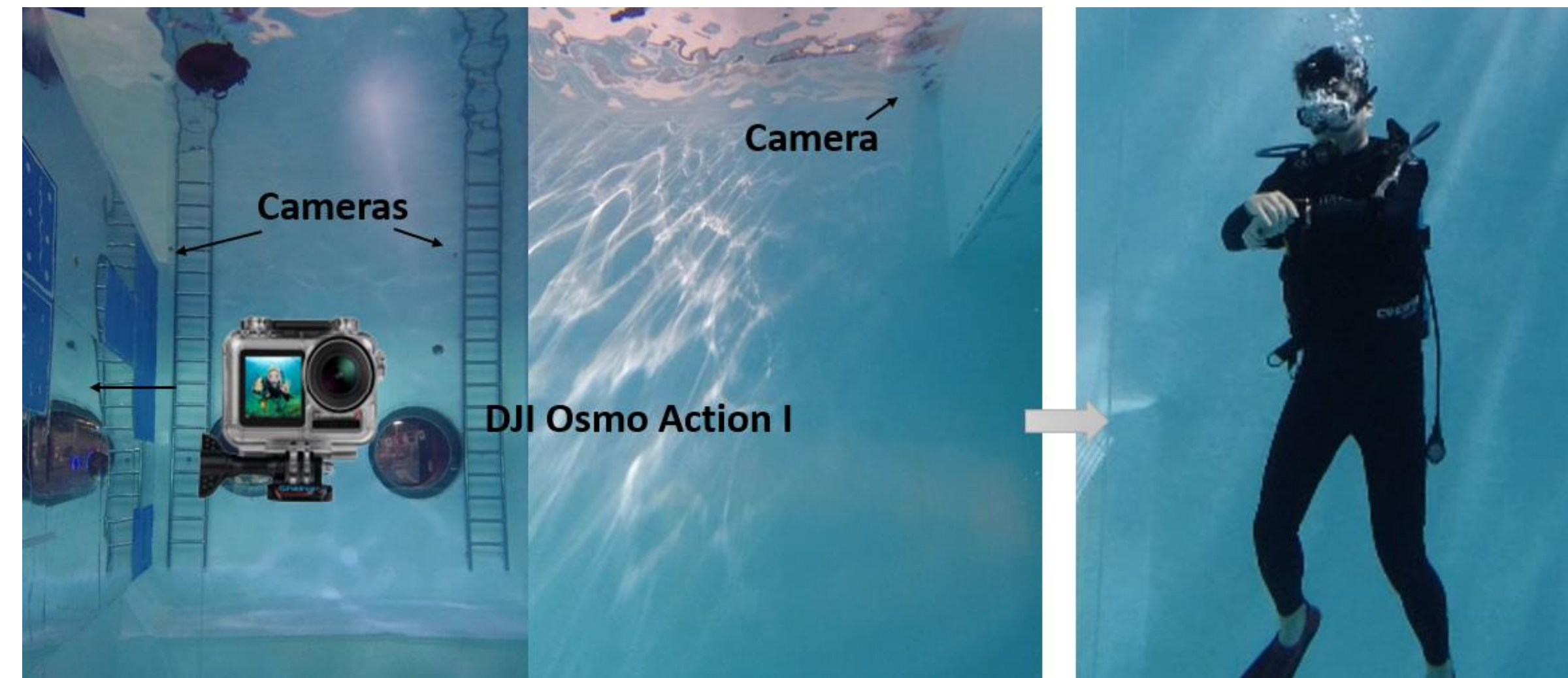
- **Device and environment:** DJI Action, in a deep diving pool in Shenzhen.
- **Bubble dataset:** 1802 labeled images, for breath state classification.

Peception of motion and breath

- **PoseNet:** Keypoints extraction with Mediapipe[3]; Post processing.
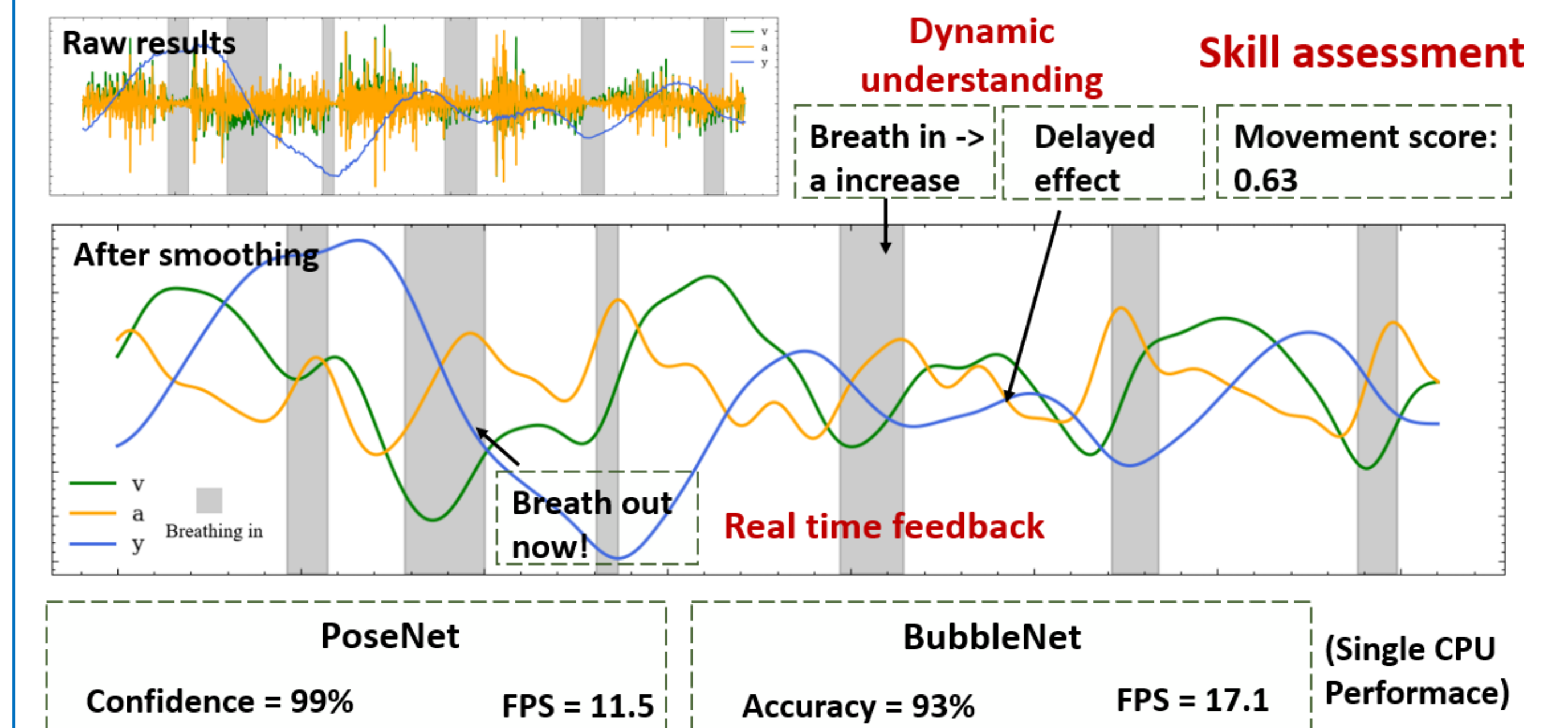- **BubbleNet:** Predict breath state from an image block around the mouth.

RL environment and diver agent training
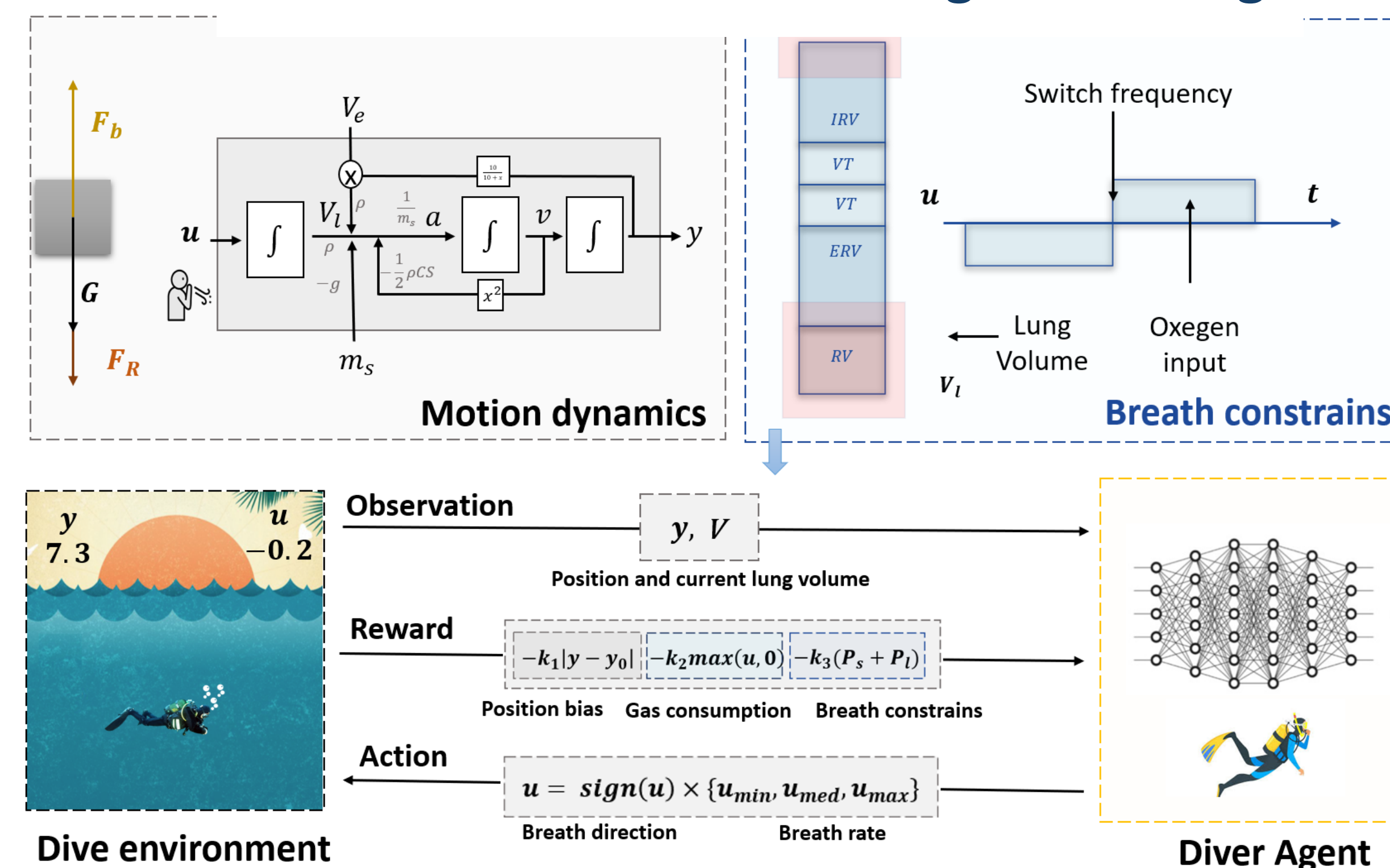
- **A Gym RL environment** based on motion dynamics and breath constrains
- **Agents trained with DQN,** with and without breath constrains
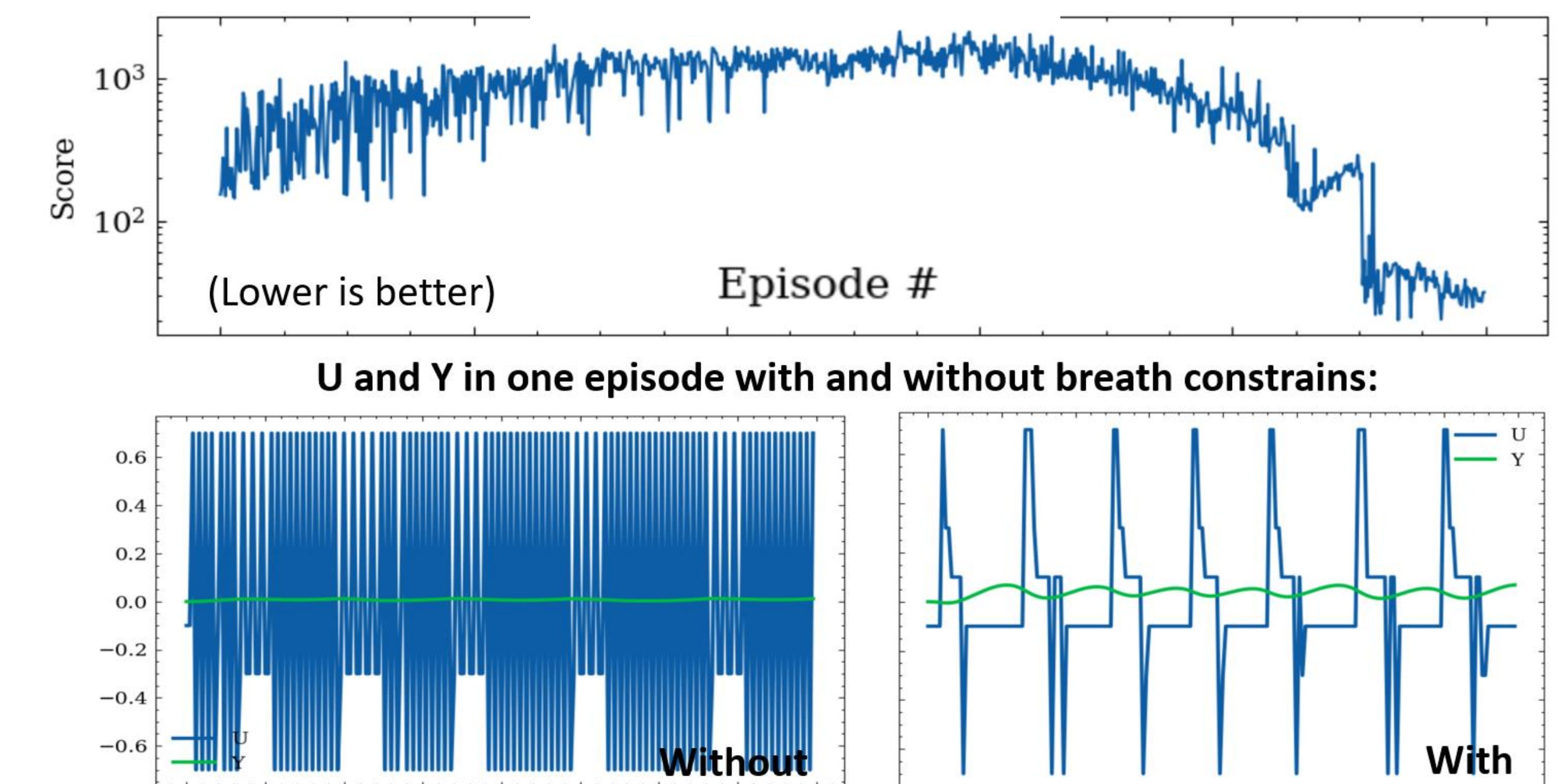
Results

Motion and breath extraction

PoseNet		BubbleNet		(Single CPU Performace)
Confidence = 99%	FPS = 11.5	Accuracy = 93%	FPS = 17.1	

- Breath and motion can **be extracted accurately in real-time even on cpu.**
- The results can be used to **understand the dynamics, provide feedback and assess diver's skill.**

Agents training

- Agent can be trained to solve the Diving environment with DQN
- Breath pattern similar to human can be obtained by adding breath constrains.

Conclusion

- PoseNet and bubbleNet can extract motion and breath states with high accuracy and speed, enabling its application on mobile device.

- A gym environment was built for the diving problem, and agents can be trained to generate the right action based on motion and breath states.

- Results demonstrated the framework's abality for assessing diver's skill, understanding the dynamics and providing real-time feedback for divers.