ĐẠI HỌC QUỐC GIA THÀNH PHỐ HỒ CHÍ MINH
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN
VNUHCM - UIT

Self Exercise Manager Service

University of Information Technology
Vietnam National University of Ho Chi Minh City - VNHCM

GROUP 6

Ly Nguyen Thuy Linh	Tran Kim Ngoc Ngan	Tran Nhat Khoa	Le Tran Quoc Khanh
22520766@gm.uit.edu.vn	22520002@gm.uit.edu.vn	22520691@gm.uit.edu.vn	22520638@gm.uit.edu.vn

I. INTRODUCTION

Nowadays students, especially Information Technology students, often lead sedentary lifestyles, which can negatively impact their health and well-being. To address this issue, we have developed a tool that encourages regular physical exercise by allowing educational institutions to monitor and evaluate students' fitness routines. This tool not only promotes healthier habits but also integrates with the academic framework to award merit points for physical activity.

II. PROBLEM STATEMENT

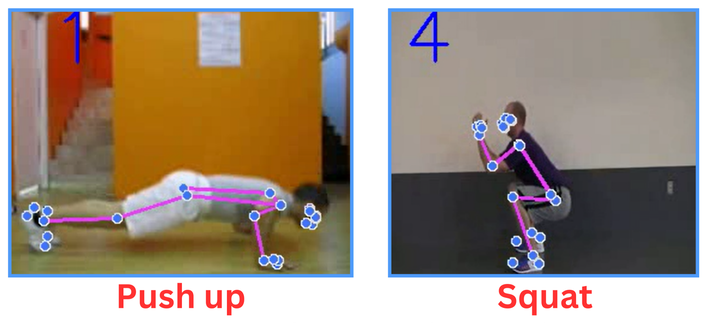
REQUIREMENTS AND CONSTRAINTS

Requirements: Maintain ~5% error margin for action counting, count valid reps, record timing, and provide fitness evaluation.
Constraints: Real-time, 480p+, one-side full-body, single person, one action type per video.

PROBLEM

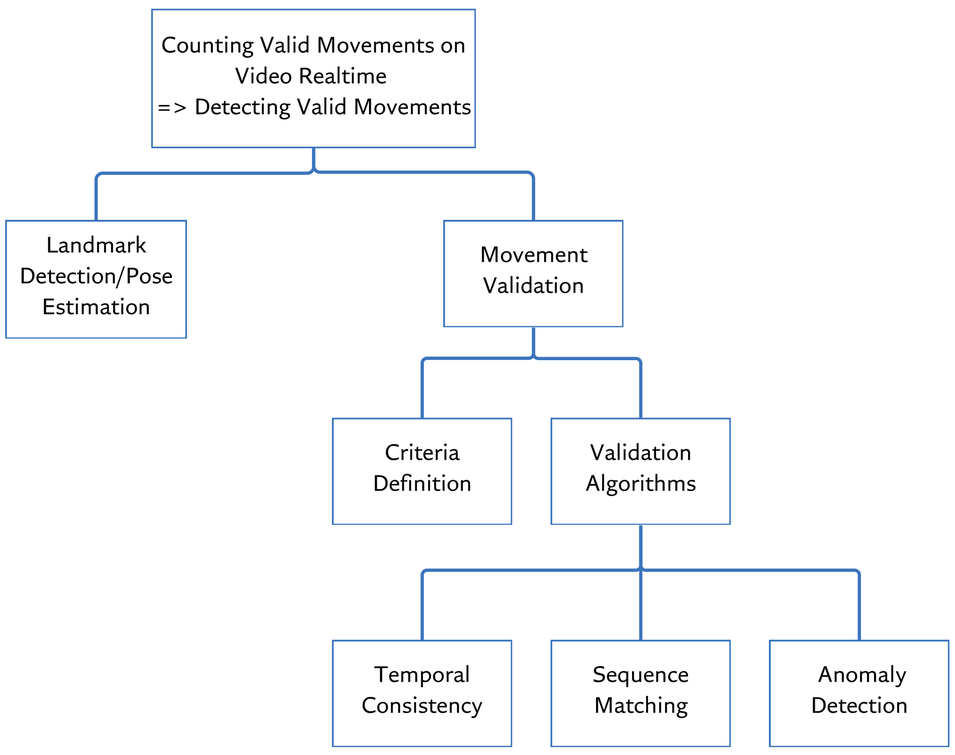
How can we effectively track, evaluate, and promote regular physical exercise within an academic environment?
To address this issue, we propose a tracking-counting actions system:

- Input: Real-time video of squat or push-ups performed by students, and the student's selected excercise (push-up or squat).
- Output: Number of valid exercise motions performed by the students, along with the duration of each motion.



Results of our solution for the problem

III. DECOMPOSITION HIERARCHY



IV. EVALUATION

DATASET

We utilize the UCF101 dataset, a comprehensive collection of videos across various action categories, including push-ups and squats. For our analysis, we select 10 videos from these two activities that meet the problem's constraints. These videos are temporarily annotated at the frame level to indicate the precise timing of each exercise action.

class	push-up	squat	total
quantity	10	10	20

```
{
    "type": "squat",
    "count": 5,
    "label": [ # [frame_id_begin, frame_id_end]
        [3, 30],
        [45, 65],
        [80, 105],
        [120, 143],
        [164, 183]
    ]
}
```

Example of Annotation Format

METRICS

Denotes

$A = [s_A, e_A]$: label interval
$B = [s_B, e_B]$: output interval
M_i, O_i : the number of label intervals and output countings of ith video
N : the number of videos

Metrics

Intersection Over Union (IOU):

$$IOU(A,B) = \begin{cases} 1, \text{if } A \in B \\ I(A,B)/U(A,B), \text{else} \end{cases}$$

Average IOU accuracy : ↑

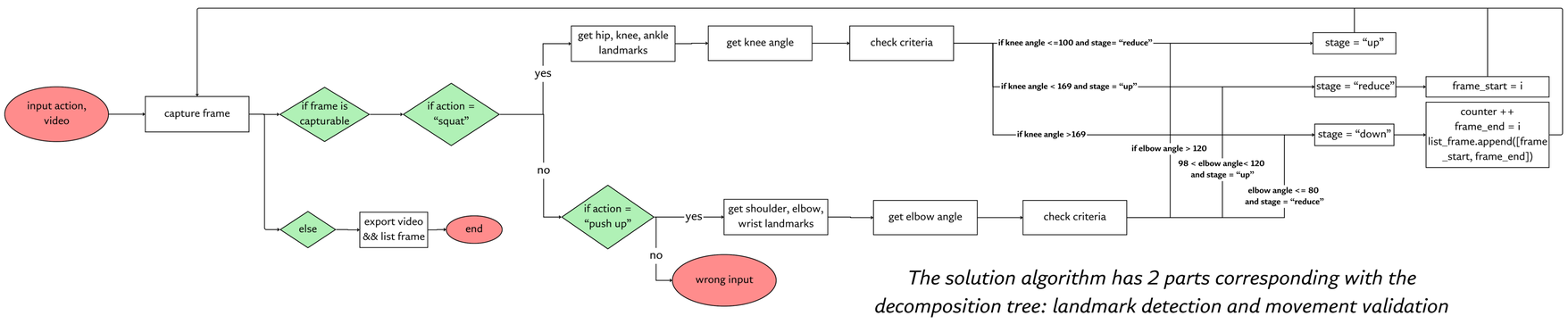
$$\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{M_i}\frac{1}{M_i}[IOU(A_{ij}, B_{ij}) > 0.9]$$

Average counting accuracy : ↑

$$\frac{1}{N}\sum_{i=1}^{N}\frac{O_i}{M_i}$$

V. METHODOLOGY AND EXPERIMENT

METHODOLOGY



The solution algorithm has 2 parts corresponding with the decomposition tree: landmark detection and movement validation

1. Landmark Detection

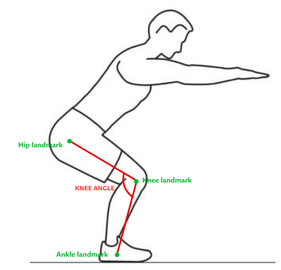
For the landmark detection problem, we employ Mediapipe, a comprehensive framework that provides essential tools for detecting and tracking landmarks in real-time. These landmarks represent key points on the body, such as joints and limbs. By capturing these points frame-by-frame, we obtain a detailed map of body positions during exercises. Utilizing the detected landmarks, we define and track specific angles (e.g., knee angle during squats, elbow angle during push-ups) over time to detect specific actions.

2. Movement Validation

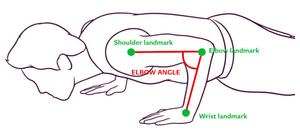
To accurately validate excercises such as push-ups and squats, we develop an algorithm with specific criteria to discriminate between different actions. The movement validation component involves an algorithm designed to measure and validate movements with high precision, minimizing false positives. Once an action is identified, we track the frames associated with it and increase the counter variable.

3. Evaluation Method

Accurate evaluation is critical for video applications, particularly for our counter system, which must not only accurately count valid actions but also capture the exact moments or processes of the movements.
We apply validation metrics using the Intersection over Union (IoU) formula to measure the accuracy of our algorithm by comparing the overlap between the detected action's start and end frames with the actual frames.
To address instances where the IoU score might inaccurately reflect performance, such as when the output frame range is entirely within the label frame range, we customize the IoU score to return a value of 1 in these cases, indicating optimal results.
Furthermore, discrepancies in the number of output frame ranges compared to label frame ranges necessitate a mapping strategy. For each range in the label frame ranges, we select the output frame range with the highest IoU score to evaluate the accuracy of our method. An IoU threshold is then employed to determine the exactness of the evaluation.



Knee angle with its landmarks



Elbow angle with its landmarks

RESULTS

The results are evaluated using Average IOU accuracy with an overlap threshold of 0.9 and Average counting accuracy for the entire dataset.

These results meet the requirement of counting error being less than 5%, and demonstrate excellent tracking performance on our annotated dataset.

	Average IOU accuracy ↑	Average Counting Accuracy ↑
Squat	0.8111	0.9688
Pushup	0.9722	0.9500

Furthermore, in real-time applications, we achieve high visual performance, as seen in the following demo:

QR to demo link