

BÁO CÁO TIẾN ĐỘ ĐỒ ÁN MÔN HỌC

MÔN: XỬ LÝ ẢNH SỐ

1. Tên đề tài

Ứng dụng Yolo trích xuất khung xương để nhận dạng hành động người trong video

2. Thành viên nhóm

- Mai Hồng Hải – MSSV: 22133014
- Nguyễn Tấn Hùng – MSSV: 22133027
- Nguyễn Ngọc Hiếu Hảo – MSSV: 22133015

3. Mục tiêu đề tài

3.1. Đặt vấn đề và Mục tiêu tổng quát

Trong bối cảnh rèn luyện sức khỏe tại nhà ngày càng phổ biến, Yoga là bộ môn được ưa chuộng nhưng lại đòi hỏi độ chính xác cao về tư thế (alignment). Việc tập sai tư thế không mang lại hiệu quả mà còn gây chấn thương xương khớp nghiêm trọng (ví dụ: đau lưng dưới khi Plank sai, chân thương đầu gối khi tập Warrior).

Do đó, nhóm em xây dựng hệ thống "Trợ lý Yoga ảo" (AI Yoga Assistant) với mục tiêu:

1. **Nhận diện tư động:** Phân biệt được người tập đang thực hiện động tác nào trong 5 tư thế phổ biến: **Downdog, Goddess, Plank, Tree, Warrior2.**
2. **Đánh giá kỹ thuật (Correction):** Phân tích hình học của khung xương để xác định tư thế đó là "**Đúng**" (**Correct**) hay "**Sai**" (**Incorrect**) dựa trên các tiêu chuẩn cơ sinh học (biomechanics).
3. **Phản hồi trực quan:** Cung cấp cảnh báo thời gian thực (Real-time feedback) để người tập điều chỉnh ngay lập tức.



Hình 1. Tree pose

Đặc điểm: Đứng trên một chân, bàn chân kia đặt lên đùi trong (hoặc bắp chân) của chân trụ, hai tay chắp trước ngực (hoặc vươn cao).



© Melina Meza

Hình 2. Plank pose

Đặc điểm: Cơ thể tạo thành một đường thẳng từ đầu đến gót chân, chống đỡ bằng khuỷu tay và ngón chân, giữ lưng thẳng.



Hình 3. Warrior2

Đặc điểm: Chân bước rộng, một chân gấp gối vuông góc, hai tay dang ngang bằng vai, mắt nhìn theo tay trước.



Hình 4. Goddess pose

Đặc điểm: Hai chân mở rộng, đầu gối gập sâu và hướng ra ngoài (giống tư thế squat), hai tay giơ lên cao vuông góc (dạng xương rồng).



Hình 5. Downdog pose

Đặc điểm: Người tập tạo thành hình dạng chữ V **ngược** với hai tay và hai chân chống xuống sàn, hông nâng cao và đẩy ra sau. Đây là một tư thế nghỉ ngơi tích cực, kéo giãn toàn bộ cơ thể.

3.2. Mục tiêu cụ thể

- **Về mặt dữ liệu:** Xử lý và chuẩn hóa bộ dữ liệu **Yoga Pose Classification** (Kaggle), khắc phục các vấn đề về nhiễu nền hoặc góc quay đa dạng.
- **Về mặt mô hình AI:**
 - Úng dụng **YOLOv8 Pose** (SOTA - State of the art) để trích xuất 17 điểm khớp (Keypoints) trên cơ thể người với tốc độ cao (đảm bảo > 20 FPS).
 - Xây dựng thuật toán phân loại (Classification) dựa trên vector đặc trưng của khung xương (thay vì dùng CNN thuần túy trên ảnh RGB để giảm tải tính toán).
- **Về mặt thuật toán đánh giá (Correction Algorithm):**
 - Xây dựng bộ thư viện hàm tính toán góc hình học (Geometric Angles) giữa các khớp.
 - Thiết lập bộ luật (Rule-based constraints) cho từng tư thế cụ thể. Ví dụ: Tư thế Plank yêu cầu 3 điểm (Vai - Hông - Cẳng chân) phải tạo thành một đường thẳng (góc xấp xỉ 180 độ).

4. Nội dung thực hiện chi tiết

4.1. Khảo sát lý thuyết và Công nghệ

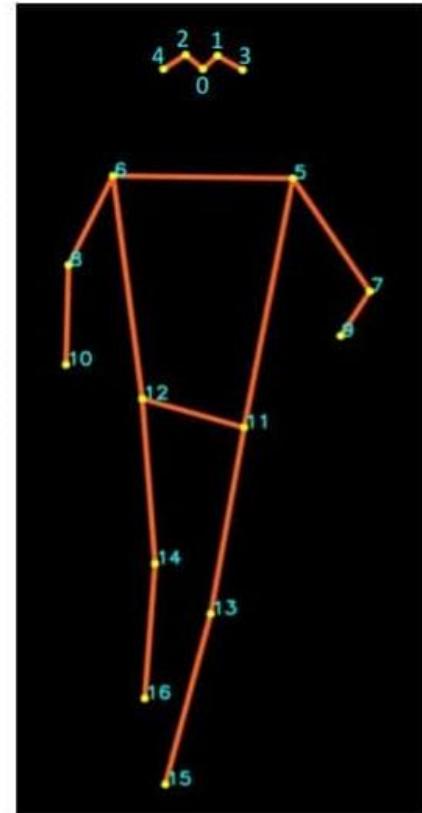
- **YOLOv8 Pose:** Tìm hiểu kiến trúc mạng nơ-ron tích chập (CNN) của YOLOv8 phiên bản Pose, cách mô hình dự đoán đồng thời Bounding Box (khung bao người) và Keypoints (các khớp). Hệ thống sử dụng định dạng COCO Keypoints gồm 17 điểm (Mũi, mắt, tai, vai, khuỷu, cổ tay, hông, gối, cổ chân...).
- **Hình học giải tích trong xử lý ảnh:**
 - Biểu diễn khung xương dưới dạng đồ thị các điểm kết nối.
 - Công thức tính góc giữa 3 điểm A, B, C (với B là khớp trung tâm) sử dụng vector:

$$\theta = \arccos \left(\frac{\mathbf{BA} \cdot \mathbf{BC}}{|\mathbf{BA}| |\mathbf{BC}|} \right)$$

- Đây là công thức cốt lõi để tính góc khuỷu tay, góc gối, độ thẳng lưng.



Index	Key point
0	Nose
1	Left-eye
2	Right-eye
3	Left-ear
4	Right-ear
5	Left-shoulder
6	Right-shoulder
7	Left-elbow
8	Right-elbow
9	Left-wrist
10	Right-wrist
11	Left-hip
12	Right-hip
13	Left-knee
14	Right-knee
15	Left-ankle
16	Right-ankle



Hình 6. Sơ đồ 17 điểm khớp chuẩn COCO Keypoints

4.2. Quy trình xử lý (System Pipeline)

Hệ thống được thiết kế theo luồng xử lý tuần tự (Sequential Pipeline):

- Input:** Video đầu vào hoặc Webcam stream.
- Pose Detection (YOLOv8):**
 - Đưa từng frame ảnh qua model.
 - Output: Tọa độ (x, y) của 17 điểm khớp và độ tin cậy (confidence score).
- Preprocessing & Normalization:**
 - Lọc bỏ các keypoint có độ tin cậy thấp (< 0.5).
 - Chuẩn hóa tọa độ về đoạn [0, 1] hoặc căn giữa theo điểm Hông (Hip center) để thuật toán không bị ảnh hưởng bởi vị trí đứng của người trong khung hình.
- Classification Module (Phân loại):**
 - Sử dụng một mô hình máy học nhẹ (như Random Forest hoặc Neural Network đơn giản) nhận input là vector tọa độ xương đã chuẩn hóa.
 - Output: Tên tư thế (Ví dụ: "Plank").
- Correction Module (Đánh giá & Sửa lỗi):**
 - Dựa vào tên tư thế nhận được, hệ thống sẽ kích hoạt bộ luật kiểm tra tương ứng.
 - Ví dụ logic kiểm tra:
 - Nếu là Plank:** Tính góc Hip_Angle (Vai-Hông-Gối). Nếu Hip_Angle < 160 độ → Cảnh báo "Hông quá cao/thấp".

- **Nếu là Tree Pose:** Tính khoảng cách Dist giữa (Cẳng chân đang co) và (Đầu gối chân trụ). Nếu Dist quá nhỏ → Cảnh báo "Không đặt chân lên khớp gối".

6. Visualizer (Hiển thị):

- Vẽ khung xương lên ảnh gốc.
- Màu khung xương: Xanh lá (Đúng) / Đỏ (Sai).
- Hiển thị thông báo text hướng dẫn sửa lỗi.

4.3. Xây dựng Bộ dữ liệu và Thực nghiệm

- **Nguồn dữ liệu:** Dataset [Yoga Pose Classification](#) gồm ~1000 ảnh cho 5 lớp.
- **Xử lý dữ liệu:**
 - Viết script chạy YOLOv8 qua toàn bộ tập ảnh để tạo ra file .csv. File này chứa các dòng dữ liệu, mỗi dòng gồm: [Class_ID, x1, y1, ..., x17, y17].
 - Dùng file .csv này để huấn luyện mô hình phân loại (Classifier).
- **Kịch bản kiểm thử (Test Scenario):**
 - Quay video thực tế các thành viên trong nhóm thực hiện 5 động tác, bao gồm cả các trường hợp **cố tình làm sai** (ví dụ: gù lưng, chùng gối) để kiểm chứng khả năng phát hiện lỗi của hệ thống.

5. Phân công nhiệm vụ chi tiết

Thành viên	MSSV	Vai trò / Nhiệm vụ chính	Chi tiết công việc
Mai Hồng Hải (Nhóm trưởng)	22133014	AI Core & System Integration Phụ trách YOLOv8 Pose và kiến trúc hệ thống	<ul style="list-style-type: none"> Nghiên cứu Ultralytics & tinh chỉnh YOLOv8 Pose (Nano/Medium) để tối ưu tốc độ–độ chính xác. Xây dựng chương trình chính bằng Python/OpenCV: đọc webcam, xử lý frame, điều phối pipeline. Tích hợp các mô-đun: Preprocessing → Classifier → Pose Correction vào hệ thống thống nhất. Thiết kế UI Overlay: hiển thị skeleton, FPS, label tư thế, cảnh báo sai tư thế. Tối ưu code, quản lý cấu trúc dự án và điều phối tiến độ nhóm.
Nguyễn Tân Hùng	22133027	Toán học & Thuật toán Đánh giá (Correction Logic)	<ul style="list-style-type: none"> Nghiên cứu tài liệu thể thao/y khoa để xác định “góc chuẩn” cho 5 tư thế Yoga (Plank, Tree, Warrior II, Goddess, Downdog). Xây dựng module GeometryUtils: tính góc, độ thẳng hàng, khoảng cách Euclidean, vector alignment. Xây dựng module PoseEvaluator: nhận keypoints đã chuẩn hóa → trả về (IsCorrect, Message).

Thành viên	MSSV	Vai trò / Nhiệm vụ chính	Chi tiết công việc
			<ul style="list-style-type: none"> Chống nhiễu: lọc rung keypoint, tránh false-alarm khi người di chuyển nhẹ. Kiểm thử và tinh chỉnh độ nhạy của các rule.
Nguyễn Ngoc Hiếu Hảo	22133015	Dữ liệu – Huấn luyện mô hình Phân loại & Testing	<ul style="list-style-type: none"> Thu thập dataset từ Kaggle + video để tăng tính đa dạng dữ liệu. Làm sạch dữ liệu: loại ảnh không có người, keypoint thiếu, tư thế không rõ ràng. Trích xuất đặc trưng: chuyển ảnh → tọa độ skeleton → vector feature chuẩn hóa. Huấn luyện mô hình phân loại tư thế (Random Forest/SVM/MLP) mục tiêu accuracy > 90%. Xây dựng bộ video kiểm thử, gán nhãn Ground Truth và đánh giá độ chính xác của hệ thống cuối. Ghi lại kết quả thí nghiệm và tạo báo cáo đánh giá.

6. Dự kiến kết quả và Hướng phát triển

6.1. Kết quả dự kiến

- Một phần mềm chạy trên PC, nhận diện được 5 tư thế Yoga với độ trễ thấp (< 100ms).
- Phát hiện chính xác các lỗi sai cơ bản: Lung cong (Plank), tay không thẳng (Warrior), chân sai vị trí (Tree).
- Báo cáo đồ án hoàn chỉnh cùng video demo thực tế.

6.2. Hướng phát triển

- Mở rộng tập dữ liệu lên 20-30 tư thế Yoga phức tạp hơn.
- Tích hợp bộ đếm thời gian (ví dụ: Đếm thời gian giữ Plank 60 giây).
- Phát triển phiên bản Mobile App để người dùng có thể tập luyện mọi lúc mọi nơi.