## DNN và CNN được thực thi ở đâu?

Mô hình DNN và CNN được thực thi trong quá trình nhận diện và phân loại hành động từ các khung xương của nhiều người. Cụ thể:

* **DNN** (Deep Neural Network) được sử dụng để học và phân loại các đặc trưng phức tạp phi tuyến tính từ dữ liệu khung xương. Quá trình dự đoán hành động thời gian thực diễn ra khi DNN nhận dữ liệu đầu vào từ các vector đặc trưng của khung xương và thực hiện phân loại hành động .
* **CNN** (Convolutional Neural Network) được thiết kế đặc biệt để xử lý dữ liệu dạng lưới như hình ảnh. Trong mô hình này, CNN có thể được sử dụng để trích xuất và nhận dạng các đặc trưng không gian từ các điểm chính trên khung xương .

Quá trình thực thi DNN và CNN diễn ra chủ yếu trên dữ liệu khung xương được trích xuất từ OpenPose, một công cụ phát hiện khung xương trong thời gian thực, cung cấp dữ liệu đầu vào cho các mô hình này để thực hiện phân loại hành động.

## lan truyền ngược và giảm chiều là gì? PCA , PAF là gì và cách nó hoạt động

**Lan truyền ngược (Backpropagation) và Giảm chiều (Dimensionality Reduction)**

* **Lan truyền ngược (Backpropagation)**: Đây là một thuật toán được sử dụng để huấn luyện mạng neural. Sau khi dự đoán được thực hiện bằng cách truyền dữ liệu qua các lớp của mạng, lan truyền ngược tính toán độ chênh lệch giữa giá trị dự đoán và giá trị thực tế thông qua hàm mất mát (Loss Function). Sau đó, thuật toán cập nhật trọng số của các kết nối giữa các neuron bằng cách sử dụng gradient descent để giảm thiểu giá trị của hàm mất mát này​(HocMayvaUngDung\_BaoCaoC…).
* **Giảm chiều (Dimensionality Reduction)**: Là quá trình giảm số lượng đặc trưng trong dữ liệu đầu vào xuống mức cần thiết, trong khi vẫn giữ lại các thông tin quan trọng. Một phương pháp phổ biến để giảm chiều là **PCA (Principal Component Analysis)**, giúp giảm số chiều dữ liệu bằng cách tìm kiếm các thành phần chính (principal components) đại diện cho thông tin quan trọng nhất trong dữ liệu​(HocMayvaUngDung\_BaoCaoC…).

**PCA và PAF**

* **PCA (Principal Component Analysis)**: PCA là một kỹ thuật thống kê được sử dụng để giảm số chiều của dữ liệu bằng cách biến đổi các biến gốc thành các biến mới (principal components) độc lập và không tương quan với nhau, giữ lại phần lớn thông tin của dữ liệu ban đầu. Trong bối cảnh xử lý khung xương người, PCA giúp giảm bớt số lượng đặc trưng mà mô hình cần xử lý, từ đó giúp tăng hiệu quả và tốc độ tính toán.
* **PAF (Part Affinity Fields)**: PAF là một thành phần của OpenPose, giúp xác định mối liên kết giữa các điểm chính của khung xương trên cơ thể người. Trong hệ thống, PAFs được sử dụng để kết nối các phần của khung xương lại với nhau, xác định mối liên hệ không gian giữa các phần cơ thể dựa trên các bản đồ liên kết (affinity maps)​(HocMayvaUngDung\_BaoCaoC…).

## vậy PCA có là thành phần của Openpose không, hay phần DNN

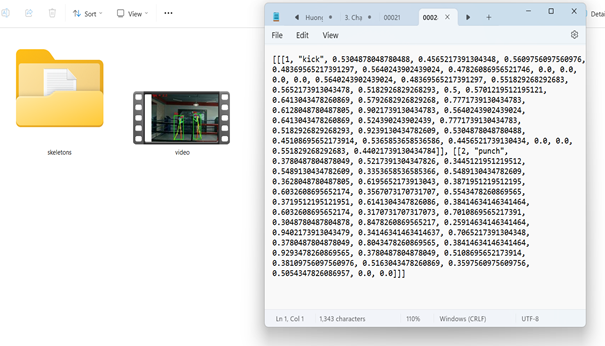
PCA (Principal Component Analysis) không phải là thành phần của OpenPose. Thay vào đó, nó là một kỹ thuật giảm chiều được sử dụng trong xử lý dữ liệu để giảm số lượng đặc trưng (features) trong khi vẫn giữ lại thông tin quan trọng.

**Vai trò của PCA và OpenPose:**

* **OpenPose**: Là một hệ thống phát hiện khung xương trong thời gian thực, dùng để xác định các điểm chính trên cơ thể người và mối liên hệ giữa chúng (PAF). OpenPose không sử dụng PCA trong việc phát hiện khung xương.
* **PCA**: Thường được áp dụng trong quá trình tiền xử lý dữ liệu hoặc trong mô hình DNN (Deep Neural Network) để giảm số chiều của dữ liệu đầu vào, giúp tối ưu hóa quá trình học và giảm độ phức tạp của mô hình.

Tóm lại, PCA không phải là thành phần của OpenPose, mà nó thường được sử dụng trong phần DNN hoặc các bước tiền xử lý dữ liệu để giảm chiều dữ liệu trước khi đưa vào mô hình học sâu.

## các thông số trong file skeleton sau khi được detect là như thế nào



Các thông số trong file skeleton mà bạn vừa đề cập đến có thể là tọa độ của các điểm chính trên khung xương của người trong ảnh hoặc video.

**Cụ thể:**

* Mỗi bộ tọa độ đại diện cho vị trí của một khớp (joint) trên cơ thể người trong không gian 2D hoặc 3D.
* Các giá trị này có thể bao gồm cả tọa độ (x, y, z) và độ tin cậy (confidence score) của mỗi khớp.

**Ví dụ về các thông số:**

* **"kick"** và **"punch"**: Đây có thể là nhãn (label) cho hành động mà hệ thống đã nhận diện từ dữ liệu khung xương.
* **Các con số**: Các dãy số có thể là tọa độ x, y, (và có thể là z nếu là dữ liệu 3D) của các điểm chính trên khung xương của người. Những con số này có thể được sử dụng làm đầu vào cho các mô hình học máy để phân loại hành động.

Nhìn vào hình ảnh bạn đã chia sẻ, có vẻ bạn đang xử lý dữ liệu khung xương từ một video, và các thông số này được sử dụng để nhận diện hành động cụ thể từ bộ dữ liệu này.

## vậy DNN của mô hình này hoạt động như thế nào , có phải sau khi openpose thực hiện công việc trích xuất các đặc trưng khung xương và gửi về DNN hay không, hãy nói chi tiết hơn cách DNN hoạt động trong mô hình này

Đúng vậy, trong mô hình này, OpenPose được sử dụng để trích xuất các đặc trưng khung xương từ video hoặc hình ảnh, sau đó những đặc trưng này được gửi đến mô hình DNN (Deep Neural Network) để thực hiện phân loại hành động.

Cách DNN hoạt động trong mô hình:

1. **Dữ liệu đầu vào**:
   * Dữ liệu đầu vào của DNN là các tọa độ khung xương (có thể là 2D hoặc 3D) mà OpenPose đã trích xuất được. Mỗi khung xương bao gồm các điểm chính đại diện cho các khớp trên cơ thể người như vai, khuỷu tay, đầu gối, v.v.
2. **Tiền xử lý dữ liệu**:
   * Các tọa độ khung xương được chuẩn hóa và có thể qua bước giảm chiều bằng các kỹ thuật như PCA (Principal Component Analysis) để giảm bớt độ phức tạp và tăng hiệu suất xử lý.
3. **Xử lý trong DNN**:
   * **Lớp đầu vào (Input Layer)**: Các tọa độ khung xương được đưa vào lớp đầu vào của DNN.
   * **Các lớp ẩn (Hidden Layers)**: Dữ liệu được truyền qua nhiều lớp ẩn của mạng neural, mỗi lớp ẩn có nhiều neuron kết nối với nhau. Tại mỗi lớp, các phép biến đổi phi tuyến tính (thông qua các hàm kích hoạt như ReLU) được áp dụng để học các đặc trưng phức tạp hơn từ dữ liệu.
   * **Lan truyền ngược (Backpropagation)**: Trong quá trình huấn luyện, thuật toán lan truyền ngược được sử dụng để điều chỉnh các trọng số (weights) của các kết nối trong mạng nhằm giảm thiểu lỗi giữa giá trị dự đoán và giá trị thực tế.
4. **Lớp đầu ra (Output Layer)**:
   * Sau khi dữ liệu đi qua các lớp ẩn, nó sẽ được đưa đến lớp đầu ra. Lớp này sẽ cho ra kết quả dự đoán hành động, thường là một xác suất cho mỗi hành động có thể xảy ra. Ví dụ, xác suất cho các hành động như "kick", "punch", "wave", v.v.
5. **Dự đoán hành động**:
   * Dựa trên xác suất từ lớp đầu ra, mô hình sẽ dự đoán hành động nào đang diễn ra trong khung hình hiện tại.

Tổng quan về quy trình:

* Quá trình này diễn ra liên tục với mỗi khung hình trong video, giúp hệ thống nhận diện hành động của một hoặc nhiều người trong thời gian thực.

DNN đóng vai trò quan trọng trong việc phân loại các hành động dựa trên thông tin khung xương, giúp hệ thống nhận diện chính xác các hành động phức tạp từ dữ liệu khung xương mà OpenPose cung cấp.

## vậy CNN ở đâu, có phải CNN là 1 phần của openpose hay không, nó hoạt động như thế nào

CNN và Vai trò của nó trong OpenPose:

**1. CNN trong OpenPose**:

* **Đúng**: CNN (Convolutional Neural Network) thực sự là một phần quan trọng của OpenPose. OpenPose sử dụng CNN để trích xuất đặc trưng từ hình ảnh đầu vào và xác định vị trí của các điểm chính trên cơ thể người (keypoints) trong không gian 2D hoặc 3D.

**2. Hoạt động của CNN trong OpenPose**:

* **Trích xuất đặc trưng (Feature Extraction)**:
  + CNN nhận đầu vào là các khung hình (images) từ video hoặc ảnh tĩnh.
  + Các lớp tích chập (convolutional layers) trong CNN trích xuất các đặc trưng từ hình ảnh, chẳng hạn như các đường viền, hình dạng và kết cấu, liên quan đến các khớp và bộ phận cơ thể.
  + Quá trình này giúp phát hiện các điểm chính (keypoints) của cơ thể, như vai, khuỷu tay, đầu gối, mắt cá chân, v.v.
* **Dự đoán vị trí điểm chính (Keypoint Detection)**:
  + Sau khi các đặc trưng được trích xuất, CNN sử dụng các bản đồ đặc trưng (feature maps) để dự đoán vị trí các điểm chính của cơ thể trong hình ảnh.
  + Một loạt các bản đồ liên kết (PAF - Part Affinity Fields) cũng được tạo ra để mô hình hóa các mối liên hệ không gian giữa các điểm chính này.
* **Lớp tích chập và pooling**:
  + CNN bao gồm các lớp tích chập (convolutional layers) để trích xuất đặc trưng, và các lớp pooling để giảm kích thước của các bản đồ đặc trưng, giữ lại thông tin quan trọng nhất.
  + Các bản đồ đặc trưng cuối cùng được sử dụng để dự đoán vị trí của các điểm chính.
* **Tính toán PAFs (Part Affinity Fields)**:
  + PAFs giúp xác định mối liên kết giữa các điểm chính trên cơ thể người trong hình ảnh. Điều này rất hữu ích khi xử lý nhiều người trong cùng một khung hình, vì nó giúp kết nối đúng các điểm chính với từng người cụ thể.

**3. Tóm tắt quy trình**:

* **CNN** trong OpenPose thực hiện việc nhận diện các điểm chính của cơ thể từ hình ảnh đầu vào.
* **DNN** trong mô hình nhận diện hành động sử dụng đầu ra từ OpenPose (dữ liệu khung xương) để phân loại và nhận diện hành động.

Tóm lại, CNN là một phần của OpenPose, đóng vai trò trong việc trích xuất đặc trưng từ hình ảnh để phát hiện các điểm chính trên cơ thể người.

## tôi muốn hỏi tiếp về hàm ReLU ở lớp ẩn DNN , nó như thế nào?

Hàm ReLU (Rectified Linear Unit) là một hàm kích hoạt rất phổ biến được sử dụng trong các lớp ẩn của mạng neural (DNN). ReLU có dạng công thức như sau:

ReLU(x)=max⁡(0,x)\text{ReLU}(x) = \max(0, x)ReLU(x)=max(0,x)

Cách hoạt động:

* **Đầu vào x**: ReLU nhận giá trị đầu vào xxx là tổng trọng số của các đầu vào trước đó và bias.
* **Đầu ra**: Nếu giá trị xxx nhỏ hơn 0, hàm ReLU sẽ trả về 0. Ngược lại, nếu xxx lớn hơn 0, nó sẽ giữ nguyên giá trị của xxx.

Ưu điểm của ReLU:

1. **Đơn giản và hiệu quả**: ReLU dễ tính toán, nhanh hơn các hàm kích hoạt phi tuyến khác như sigmoid hoặc tanh.
2. **Giải quyết vấn đề vanishing gradient**: Do giá trị đầu ra của ReLU cho các giá trị dương không bị giới hạn, nó giúp giảm thiểu vấn đề vanishing gradient thường gặp với các hàm kích hoạt khác, nơi gradient có thể trở nên rất nhỏ và làm chậm quá trình học.

Vai trò trong DNN:

* Trong các lớp ẩn của DNN, ReLU được áp dụng sau mỗi phép tính tổng trọng số, giúp mạng học được các đặc trưng phi tuyến từ dữ liệu.
* Bằng cách sử dụng ReLU, DNN có thể học các mối quan hệ phức tạp trong dữ liệu, từ đó cải thiện khả năng phân loại và dự đoán.

Hạn chế:

* **Dead Neurons**: Nếu một neuron liên tục nhận giá trị âm, nó sẽ không cập nhật trong quá trình học vì đầu ra của nó luôn là 0. Điều này dẫn đến vấn đề gọi là "dead neurons".

Hàm ReLU là một công cụ mạnh mẽ trong việc giúp DNN học các đặc trưng phức tạp từ dữ liệu một cách nhanh chóng và hiệu quả.