

# Ước tính tư thế con người AI: Phát hiện tư thế Yoga và Điều chỉnh

Rutuja Gajbhiye, Snehal Jarag, Pooja Gaikwad, Shweta Koparde

aGiáo sư sinh viên, Khoa Kỹ thuật Máy tính, Khoa Kỹ thuật Máy tính, Trường Cao đẳng Kỹ thuật và Nghiên cứu Pimpri Chinchwad, 412101, Ravet, Pune, Maharashtra.

Tóm tắt:- Điều quan trọng nhất của các tư thế yoga đã được biết đến trên toàn thế giới và chứng minh những lợi ích sức khỏe mà các nhà hiền triết cổ đại đã thuyết giảng. Khi yoga trở nên quan trọng hơn, yoga phải đối mặt với những thách thức quan trọng sau: Công nghệ thị giác máy tính cung cấp một giải pháp đầy hứa hẹn để đánh giá tư thế con người. Tuy nhiên, những kỹ thuật này hiếm khi được sử dụng trong các lĩnh vực sức khỏe và tập thể dục, và không có tài liệu tham khảo hoặc dự án cụ thể nào. Được đặt theo tên của yoga. Sách trắng này mô tả các công nghệ khác nhau có thể được sử dụng để ước tính tư thế và tóm tắt những cách tốt nhất để sử dụng chúng dựa trên mức độ dễ sử dụng của ứng dụng ứng dụng Android của bạn. Phần sau đây mô tả phương pháp được sử dụng để cung cấp ước tính tư thế yoga trong các ứng dụng Android, cách ứng dụng được lập mô hình và cách hoạt động của từng thành phần. Ước tính tư thế là một nhánh của thị giác máy tính liên quan đến việc nhận dạng các bộ phận riêng lẻ tạo nên cơ thể (thường là cơ thể người). Có một số cách để đạt được điều này. Cách tiếp cận tôi sử dụng bắt đầu bằng việc chuyển hình ảnh đến thông qua bộ phân loại CNN được đào tạo để tìm kiếm mọi người. Khi các tư thế cơ thể người được nhận dạng, mạng ước tính tư thế sẽ tìm kiếm các khớp và chi đã được luyện tập. Sau đó, máy tính có thể hiển thị hình ảnh cho người dùng bằng cách sử dụng các điểm đánh dấu xác định các bộ phận của cơ thể.

Từ khóa:- Deep Learning, Machine Learning, Pose Estimation, OpenPose, PostNet, YogaPoses, CNN.

101. GIỚI THIỆU

Ước tính tư thế con người là một kỹ thuật thị giác máy tính được sử dụng để dự đoán vị trí/tư thế hoặc vị trí khớp của một bộ phận cơ thể người. Để thực hiện tư thế này, các khớp hay còn gọi là các điểm chính của cơ thể người, chẳng hạn như cổ tay, khuỷu tay, đầu gối và mắt cá chân, được xác định trong hình ảnh hoặc video. Khi hình ảnh được nhập vào mô hình ước tính tư thế, tọa độ của các bộ phận cơ thể được nhận dạng này được xác định là đầu ra với giá trị độ tin cậy cho biết độ chắc chắn của cơ thể người.

Trong thời hiện đại ngày nay, các kỹ thuật học máy và học sâu đã được chứng minh là rất quan trọng đối với các nhiệm vụ khám phá đối tượng. Chúng tôi có thể sử dụng mô hình một cách hiệu quả để nhận biết các bộ phận cơ thể quan trọng khác nhau và ước tính tư thế của người dùng trong thời gian thực. Có nhiều phương pháp khác nhau mà chúng ta có thể sử dụng để ước tính tư thế. Bài viết này mô tả sự phát triển của ước tính tư thế con người trong những năm qua và cách Postnet phù hợp nhất cho dự án của chúng tôi (ước tính tư thế Yoga trên Android). Dự án này cung cấp ước tính tư thế theo thời gian thực cho phía khách hàng

ước tính và chỉnh sửa tư thế yoga. Ước tính thời gian thực có thể được thực hiện với Tensorflow. Những điểm này được sử dụng để vẽ khung xương của tư thế con người, từ đó xác định góc giữa các điểm này. Điều này cho phép bạn sửa đổi các tư thế yoga của mình một cách hiệu quả. Phương pháp này được sử dụng trong các ứng dụng Android cùng với các mô-đun Chuyển văn bản thành giọng nói và Giọng nói thành văn bản của Google để cho phép người dùng thực hành yoga rất hiệu quả. Dù sao, các vấn đề về yoga, giống như một số bài tập khác, điều quan trọng nhất là phải thực hành cẩn thận, vì tư thế sai trong buổi tập yoga sẽ không hiệu quả và có thể gây khó chịu cho mọi người.

Ước tính của con người là một vấn đề khó giải quyết trong lĩnh vực thị giác máy tính. Nó liên quan đến việc xác định vị trí các khớp của con người trong một hình ảnh hoặc video để tạo thành một đại diện cho bộ xương. Tự động phát hiện tư thế của một người trong ảnh là một nhiệm vụ khó khăn vì nó phụ thuộc vào nhiều khía cạnh, bao gồm B. Tỷ lệ và độ phân giải của ảnh, thay đổi ánh sáng, tiếng ồn xung quanh, thay đổi quần áo, môi trường và tương tác giữa người với môi trường. Ứng dụng ước tính tư thế con người đã thu hút nhiều nhà nghiên cứu trong lĩnh vực này là tập thể dục và thể dục. Đây là một thực hành cổ xưa bắt đầu ở Ấn Độ nhưng hiện đã nổi tiếng thế giới vì nhiều lợi ích về tinh thần, thể chất và tinh thần. Tuy nhiên, vấn đề với yoga là, giống như bất kỳ bài tập nào khác, tư thế không đúng trong buổi tập yoga có thể không hiệu quả và có khả năng gây hại.

Do đó, bạn cần một người hướng dẫn để giám sát phiên và điều chỉnh tư thế cá nhân của bạn. Tất cả người dùng không có quyền truy cập hoặc tài nguyên cho người hướng dẫn, vì điều này, chúng tôi có thể sử dụng ứng dụng dựa trên trí tuệ nhân tạo để xác định các tư thế yoga và cung cấp phản hồi được cá nhân hóa để cải thiện các thay đổi cá nhân của bạn. Trong những năm gần đây, việc ước tính tư thế con người đã được hưởng lợi từ việc học sâu, với những cải tiến đáng kể về hiệu suất. Phương pháp học máy cung cấp một cách dễ dàng hơn để ánh xạ cấu trúc hơn là xử lý sự phụ thuộc giữa các cấu trúc. Bằng cách sử dụng máy học và học sâu theo cách thủ công, chúng tôi đã xác định được 5 tư thế tập thể dục là kéo xà, gặp người với bóng kiểu Thụy Sĩ, chống đẩy, đập xe và đi bộ.

Một trong những đánh giá thu hút nhiều nhà phân tích trong lĩnh vực này, Sử dụng, là Thực hành và Sức khỏe. Loại hình hoạt động đa dạng là Yoga, một môn tập luyện bắt nguồn từ Ấn Độ nhưng giờ đây 444 đang tôn vinh toàn bộ Lợi ích. Ước tính tư thế từ video đóng một vai trò quan trọng trong việc phủ nội dung kỹ thuật số và thông tin về thể giới thực trong thực tế tăng cường, cho phép nhận dạng ngôn ngữ ký hiệu và kiểm soát cử chỉ toàn cơ thể. Ước tính tư thế trong các ứng dụng thể dục cho yoga, khiêu vũ và trị liệu thể dục bao gồm các tư thế khác nhau (ví dụ: hàng trăm tư thế yoga), nhiều mức độ tự do, khớp cần (ví dụ: cơ thể và các tư thế khác).

đối tượng che khuất các chi khỏi tầm nhìn của máy ảnh), và đặc biệt vì một số lý do.

II. CÔNG VIỆC CÓ LIÊN QUAN

Chúng tôi xác định một số phương pháp tiên tiến nhất để ước tính tư thế giúp ước tính chính xác các tư thế của con người dưới nhiều cấu hình cảm biến, số lần chụp và số lượng cá nhân trên mỗi lần chụp. Toshev và cộng sự. [7] là những người đầu tiên sử dụng mạng nơ-ron sâu để cải thiện khả năng phát hiện tư thế, tìm vị trí của từng khớp cơ thể bằng hồi quy trên các tính năng của CNN. Newel et al. [4] giới thiệu một kiến trúc mạng lưới thần kinh đồng hồ cát xếp chồng lên nhau, sử dụng quy trình xử lý lặp đi lặp lại từ dưới lên và từ trên xuống để đạt được các dự đoán chính xác về tư thế đơn lẻ. Ngụy và cộng sự. [8] đề xuất một kiến trúc khác, sử dụng nhiều mạng tích chập để tinh chỉnh các ước tính chung qua các lần truyền tuần tự. Thay vì dữ liệu máy ảnh RGB, Shotton et al. [6] sử dụng các bản đồ độ sâu đơn do Microsoft Kinect chụp để dự đoán vị trí 3D của các khớp thông qua phương pháp nhận dạng đối tượng. Bogo et al. [2] ước tính tư thế 3D, cũng như hình dạng lưới 3D, chỉ sử dụng các hình ảnh RGB đơn lẻ.

Một lĩnh vực nghiên cứu quan trọng cũng tập trung vào việc phát hiện tư thế của nhiều người trong một lần chụp. Papandreou et al. [5] phát hiện nhiều tư thế thông qua quy trình hai giai đoạn, đầu tiên xác định các hộp giới hạn có thể có cho mọi người, sau đó phát hiện các điểm chính của tư thế trong mỗi hộp giới hạn. Ngược lại, Cao et al. [3] sử dụng Trường sở thích một phần để ước tính tư thế của nhiều người trong một cảnh trong thời gian thực mà không cần xác định từng người trước. Cao et al. [3] đã mã nguồn mở công việc của họ dưới dạng một dự án có tên là OpenPose, mà chúng tôi sử dụng cho Pose Trainer.

Ước tính tư thế cho phép chúng tôi phân tích tư thế tĩnh của con người, điều này sẽ cung cấp thông tin có giá trị liên quan đến tính đúng đắn của tư thế. Zell và cộng sự. [9] sử dụng một cách tiếp cận thú vị để phân tích các chuyển động vật lý, trong đó cơ thể được biểu diễn như một hệ thống lò xo khối lượng và được sử dụng để tìm các lực và mô-men xoắn truyền qua các khớp của cơ thể. Chúng tôi nhận thấy rằng, bằng cách sử dụng các thông số kỹ thuật của bài tập và phản hồi từ các chuyên gia, chúng tôi có thể thực hiện một cách tiếp cận đơn giản hơn để phân tích vật lý, phân tích các góc và khoảng cách giữa các điểm chính chung để cung cấp phản hồi quan trọng cho người dùng mà không cần mô phỏng vật lý đầy đủ.

Chúng tôi xác định một số phương pháp tiên tiến nhất để ước tính tư thế giúp ước tính chính xác các tư thế của con người dưới nhiều cấu hình cảm biến, số lần chụp và số lượng cá nhân trên mỗi lần chụp. Toshev và cộng sự. [7] là những người đầu tiên sử dụng mạng nơ-ron sâu để cải thiện khả năng phát hiện tư thế, tìm vị trí của từng khớp cơ thể bằng hồi quy trên các tính năng của CNN. Newel et al. [4] giới thiệu một kiến trúc mạng lưới thần kinh đồng hồ cát xếp chồng lên nhau, sử dụng quy trình xử lý lặp đi lặp lại từ dưới lên và từ trên xuống để đạt được các dự đoán chính xác về tư thế đơn lẻ. Ngụy và cộng sự. [8] đề xuất một kiến trúc khác, sử dụng nhiều mạng tích chập để tinh chỉnh các ước tính chung qua các lần truyền tuần tự. Thay vì dữ liệu máy ảnh RGB, Shotton et al. [6] sử dụng các bản đồ độ sâu đơn do Microsoft Kinect chụp để dự đoán vị trí 3D của các khớp thông qua phương pháp nhận dạng đối tượng. Bogo et al. [2] ước tính tư thế 3D, cũng như hình dạng lưới 3D, chỉ sử dụng các hình ảnh RGB đơn lẻ.

Một lĩnh vực nghiên cứu quan trọng cũng tập trung vào việc phát hiện tư thế của nhiều người trong một lần chụp. Papandreou et al. [5] phát hiện nhiều tư thế thông qua quy trình hai giai đoạn, đầu tiên xác định các hộp giới hạn có thể có cho mọi người, sau đó phát hiện các điểm chính của tư thế trong mỗi hộp giới hạn. Ngược lại, Cao et al. [3] sử dụng Trường sở thích một phần để ước tính tư thế của nhiều người trong một cảnh trong thời gian thực mà không cần xác định từng người trước. Cao et al. [3] đã mã nguồn mở công việc của họ dưới dạng một dự án có tên là OpenPose, mà chúng tôi sử dụng cho Pose Trainer.

Ước tính tư thế cho phép chúng tôi phân tích tư thế tĩnh của con người, điều này sẽ cung cấp thông tin có giá trị liên quan đến tính đúng đắn của tư thế. Zell và cộng sự. [9] sử dụng một cách tiếp cận thú vị để phân tích các chuyển động vật lý, trong đó cơ thể được biểu diễn như một hệ thống lò xo khối lượng và được sử dụng để tìm các lực và mô-men xoắn truyền qua các khớp của cơ thể. Chúng tôi nhận thấy rằng, bằng cách sử dụng các thông số kỹ thuật của bài tập và phản hồi từ các chuyên gia, chúng tôi có thể thực hiện một cách tiếp cận đơn giản hơn để phân tích vật lý, phân tích các góc và khoảng cách giữa các điểm chính chung để cung cấp phản hồi quan trọng cho người dùng mà không cần mô phỏng vật lý đầy đủ.

Chúng tôi xác định một số phương pháp tiên tiến nhất để ước tính tư thế giúp ước tính chính xác các tư thế của con người dưới nhiều cấu hình cảm biến, số lần chụp và số lượng cá nhân trên mỗi lần chụp. Toshev và cộng sự. [7] là những người đầu tiên sử dụng mạng nơ-ron sâu để cải thiện khả năng phát hiện tư thế, tìm vị trí của từng khớp cơ thể bằng cách sử dụng hồi quy trên các tính năng của CNN. Newel et al. [4] giới thiệu một kiến trúc mạng lưới thần kinh đồng hồ cát xếp chồng lên nhau, sử dụng quy trình xử lý lặp đi lặp lại từ dưới lên và từ trên xuống để đạt được các dự đoán chính xác về tư thế đơn lẻ. Ngụy và cộng sự. [8] đề xuất một kiến trúc khác, sử dụng nhiều mạng tích chập để tinh chỉnh các ước tính chung qua các lần truyền tuần tự. Thay vì dữ liệu máy ảnh RGB, Shotton et al. [6] sử dụng các bản đồ độ sâu đơn do Microsoft Kinect chụp để dự đoán vị trí 3D của các khớp thông qua phương pháp nhận dạng đối tượng. Bogo et al. [2] ước tính tư thế 3D, cũng như hình dạng lưới 3D, chỉ sử dụng các hình ảnh RGB đơn lẻ.

Một lĩnh vực nghiên cứu quan trọng cũng tập trung vào việc phát hiện tư thế của nhiều người trong một lần chụp. Papandreou et al. [5] phát hiện nhiều tư thế thông qua quy trình hai giai đoạn, đầu tiên xác định các hộp giới hạn có thể có cho mọi người, sau đó phát hiện các điểm chính của tư thế trong mỗi hộp giới hạn. Ngược lại, Cao et al. [3] sử dụng Trường sở thích một phần để ước tính tư thế của nhiều người trong một cảnh trong thời gian thực mà không cần xác định từng người trước. Cao et al. [3] đã mã nguồn mở công việc của họ dưới dạng một dự án có tên là OpenPose, mà chúng tôi sử dụng cho Pose Trainer.

Ước tính tư thế cho phép chúng tôi phân tích tư thế tĩnh của con người, điều này sẽ cung cấp thông tin có giá trị liên quan đến tính đúng đắn của tư thế. Zell và cộng sự. [9] sử dụng một cách tiếp cận thú vị để phân tích các chuyển động vật lý, trong đó cơ thể được biểu diễn như một hệ thống lò xo khối lượng và được sử dụng để tìm các lực và mô-men xoắn truyền qua các khớp của cơ thể. Chúng tôi nhận thấy rằng, bằng cách sử dụng các thông số kỹ thuật của bài tập và phản hồi từ các chuyên gia, chúng tôi có thể thực hiện một cách tiếp cận đơn giản hơn để phân tích vật lý, phân tích các góc và khoảng cách.

giữa các điểm chính chung để cung cấp phản hồi quan trọng cho người dùng mà không cần mô phỏng vật lý đầy đủ.

Chúng tôi xác định một số phương pháp tiên tiến nhất để ước tính tư thế giúp ước tính chính xác các tư thế của con người dưới nhiều cấu hình cảm biến, số lần chụp và số lượng cá nhân trên mỗi lần chụp. Toshev và cộng sự. [7] là những người đầu tiên sử dụng mạng nơ-ron sâu để cải thiện khả năng phát hiện tư thế, tìm vị trí của từng khớp cơ thể bằng hồi quy trên các tính năng của CNN. Newell et al. [4] giới thiệu một kiến trúc mạng lưới thần kinh đồng hồ cát xếp chồng lên nhau, sử dụng quy trình xử lý lặp đi lặp lại từ dưới lên và từ trên xuống để đạt được các dự đoán chính xác về tư thế đơn lẻ. Ngụy và cộng sự. [8] đề xuất một kiến trúc khác, sử dụng nhiều mạng tích chập để tính chính các ước tính chung qua các lần truyền tuần tự. Thay vì dữ liệu máy ảnh RGB, Shotton et al. [6] sử dụng các bản đồ độ sâu đơn do Microsoft Kinect chụp để dự đoán vị trí 3D của các khớp thông qua phương pháp nhận dạng đối tượng. Bogo et al. [2] ước tính tư thế 3D, cũng như hình dạng lưới 3D, chỉ sử dụng các hình ảnh RGB đơn lẻ.

Một lĩnh vực nghiên cứu quan trọng cũng tập trung vào việc phát hiện tư thế của nhiều người trong một lần chụp. Papandreou et al. [5] phát hiện nhiều tư thế thông qua quy trình hai giai đoạn, đầu tiên xác định các hộp giới hạn có thể có cho mọi người, sau đó phát hiện các điểm chính của tư thế trong mỗi hộp giới hạn. Ngược lại, Cao et al. [3] sử dụng Trường sở thích một phần để ước tính tư thế của nhiều người trong một cảnh trong thời gian thực mà không cần xác định từng người trước. Cao et al. [3] đã mã nguồn mở công việc của họ dưới dạng một dự án có tên là OpenPose, mà chúng tôi sử dụng cho Pose Trainer.

Ước tính tư thế cho phép chúng tôi phân tích tư thế tĩnh của con người, điều này sẽ cung cấp thông tin có giá trị liên quan đến tính đúng đắn của tư thế. Zell và cộng sự. [9] sử dụng một cách tiếp cận thú vị để phân tích các chuyển động vật lý, trong đó cơ thể được biểu diễn như một hệ thống lò xo khối lượng và được sử dụng để tìm các lực và mô-men xoắn truyền qua các khớp của cơ thể. Chúng tôi nhận thấy rằng, bằng cách sử dụng các thông số kỹ thuật của bài tập và phản hồi từ các chuyên gia, chúng tôi có thể thực hiện một cách tiếp cận đơn giản hơn để phân tích vật lý, phân tích các góc và khoảng cách giữa các điểm chính chung để cung cấp phản hồi quan trọng cho người dùng mà không cần mô phỏng vật lý đầy đủ.

Chúng tôi đã xác định một số phương pháp ước tính tư thế con người tiên tiến giúp ước tính chính xác tư thế con người theo các cấu hình cảm biến, độ phơi sáng và số lượng đối tượng trên mỗi lần phơi sáng khác nhau. Họ là những người đầu tiên sử dụng mạng lưới thần kinh sâu để cải thiện khả năng phát hiện tư thế và sử dụng hồi quy chức năng CNN để xác định vị trí các khớp trong mỗi cơ thể. Họ đã giới thiệu một kiến trúc mạng nơ-ron sử dụng các phép lặp từ dưới lên và từ trên xuống để đạt được các dự đoán tư thế đơn chính xác. Họ đề xuất một kiến trúc sử dụng nhiều mạng tích chập để cải thiện ước tính chung trong các đường dẫn tuần tự. Thay vì sử dụng dữ liệu máy ảnh RGB, một bản đồ độ sâu duy nhất do Microsoft Kinect chụp được sử dụng để dự đoán vị trí 3D của khớp thông qua phương pháp phát hiện đối tượng. Ước tính tư thế 3D với hình ảnh RGB.

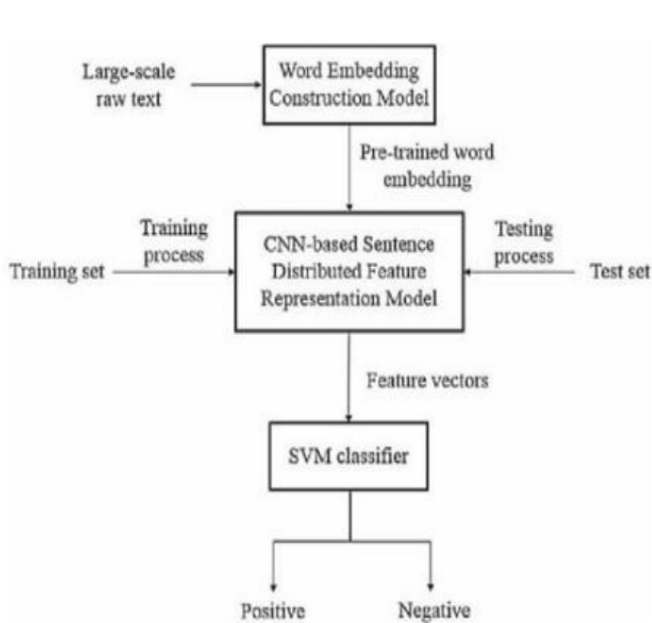
Một số nghiên cứu phổ biến cũng tập trung vào việc phát hiện tư thế của một số người trong một lần chụp. Họ xác định một số tư thế thông qua quy trình hai cấp độ, trước tiên tìm ra các giới hạn khả thi cho con người, sau đó phát hiện các điểm chính của tư thế trong mỗi hộp giới hạn. Ngược lại, sử dụng Part Affinity Fields để ước tính tư thế của một vài người trong một cảnh trong thời gian thực mà không cần nhận thức các cá nhân nam hoặc nữ trước. Họ đã mã nguồn mở các bức tranh của mình như một dự án được gọi là OpenPose, mà chúng tôi sử dụng cho Pose Trainer. Ước tính tư thế con người nghiên cứu tư thế tĩnh của con người, với mục đích đưa ra những sự thật quý giá liên quan đến tính đúng đắn của tư thế. Chúng tôi sử dụng một kỹ thuật thú vị để đánh giá các chuyển động của cơ thể, trong đó khung được thể hiện dưới dạng một thiết bị lò xo khối lượng và được sử dụng để xác định các lực và mô-men xoắn.

Họ đã quan sát thấy rằng với việc sử dụng các bài tập từ các chuyên gia, chúng tôi sẽ áp dụng một kỹ thuật ít khó khăn hơn để đánh giá cơ thể, đọc các góc và khoảng cách giữa các điểm chính chung để đưa ra những nhận xét quan trọng cho khách hàng trong khi không có mô phỏng cơ thể hoàn chỉnh.

### III. PHƯƠNG PHÁP NGHIÊN CỨU

Phần này nêu chi tiết phương pháp được sử dụng để xác định mô hình (Hình 1). Giao diện người dùng cho phép người dùng tìm hiểu về các tư thế yoga mà hệ thống cung cấp và lợi ích của chúng. Người dùng có thể thực hiện hoặc thực hành từng tư thế yoga một lần. Cả hướng dẫn bằng văn bản và bằng lời nói đều được cung cấp cho người dùng để sửa tư thế. Nguồn cấp dữ liệu video trực tiếp từ webcam được sử dụng để ghi lại chuyển động của người dùng. PostNet được sử dụng để phát hiện và sửa các tư thế yoga. PostNet là một mô hình học máy cung cấp một hệ thống phát hiện tư thế con người liên tục. Hệ thống này cho phép người dùng lần đầu tiên cố định vị trí của họ trước máy ảnh. Người dùng có thể thực hiện tư thế yoga như trong hình. Đây là nơi ghi lại các tư thế yoga của người dùng. Các điểm chính được xác định được vẽ trên canvas video. Những điểm chính này được sử dụng để so sánh tư thế của người dùng với tư thế yoga mục tiêu để xem có cần chỉnh sửa gì không. Nếu hai tư thế có trạng thái tương đồng cao thì tư thế của người dùng được coi là hoàn hảo.

Nếu tư thế yoga của người dùng không khớp với tọa độ của tư thế yoga mục tiêu, hệ thống sẽ tạo hướng dẫn để người dùng sửa đổi tư thế của họ. Người dùng có thể làm theo hướng dẫn của người hướng dẫn và sửa lỗi. Hướng dẫn bằng văn bản được cung cấp làm đầu vào cho API tổng hợp giọng nói JavaScript, cung cấp cho người dùng hướng dẫn bằng lời nói để chỉnh sửa tư thế. Sau khi tập xong yoga, người dùng có thể tiếp tục buổi tập hoặc kết thúc buổi tập.



Hình 1: Phương pháp nghiên cứu chi tiết

A. Thu thập tập dữ liệu:

Bộ dữ liệu được sử dụng cho dự án này là một phần của bộ sưu tập mã nguồn mở và được cung cấp công khai. Các tư thế yoga được sử dụng trong dự án là Padmasana (Tư thế hoa sen), Shavasana (Tư thế xác chết), Tadasana (Tư thế ngọn núi), Trikonasana (Tư thế tam giác) và Vrikshasana (Tư thế cái cây) là những tư thế yoga khác nhau. Tốc độ quay video là 30 FPS (khung hình/giây). Các video phải được quay trong nhà cách máy ảnh 4 mét. Chúng tôi đã tạo một bộ dữ liệu mà tất cả các đối tượng đều có thể thực hiện các tư thế yoga và sử dụng để xây dựng một hệ thống nhận dạng tư thế yoga mạnh mẽ. Độ dài trung bình của video phải là 45-60 giây. Các tư thế yoga khác nhau được thực hiện trong các khung Video được hiển thị trong các video có chủ đề khác nhau được sử dụng cho các bộ đào tạo, kiểm tra và xác nhận.

B. Tiền xử lý dữ liệu:

Bước đầu tiên trong quá trình tiền xử lý dữ liệu là sử dụng thư viện OpenPose để trích xuất các điểm chính của tư thế trong khung hình video. Trích xuất tư thế được thực hiện ngoại tuyến và trực tuyến trong thời gian thực và các điểm chính được xác định từ đầu vào máy ảnh được gửi đến mô hình. OpenPose chạy từng khung hình của video và đầu ra tương ứng cho từng khung hình được lưu ở định dạng JSON. Dữ liệu JSON này chứa vị trí của bộ phận cơ thể của mỗi người được xác định trong hình ảnh video. Chúng tôi đã sử dụng cài đặt mặc định để trích xuất các điểm quan trọng của tư thế để có hiệu suất lý tưởng.

Dự án bao gồm dữ liệu JSON được truy xuất và lưu trữ trong một mảng NumPy theo trình tự 45 khung. 60% tập dữ liệu được sử dụng để đào tạo, 20 phần trăm được sử dụng để thử nghiệm và 20 phần trăm được sử dụng để xác thực. Dữ liệu đào tạo chứa 45 khung gồm 7989 chuỗi, mỗi chuỗi chứa tọa độ 2D của 18 điểm chính được chụp bởi OpenPose. Dữ liệu xác thực bao gồm 84 trình tự như vậy và dữ liệu thử nghiệm chứa 84 trình tự.

IV. CHỈ SỐ ĐÁNH GIÁ

A. Điểm phân loại

Điểm phân loại đề cập đến những gì thường được coi là độ chính xác của mô hình. Điều này có thể được giải thích là tỷ lệ phần trăm của tổng số dự đoán đúng trong tổng số mẫu đầu vào.

$$\text{Accuracy} = \frac{\text{Number of correct predictions}}{\text{Total number of predictions}}$$

Trong trường hợp phân loại nhiều lớp, phép đo này cho kết quả tốt khi số lượng mẫu trong mỗi lớp gần như giống nhau.

B. Ma trận nhầm lẫn

Ma trận nhầm lẫn đại diện cho một ma trận mô tả hoàn hảo độ chính xác của mô hình. Có bốn thuật ngữ quan trọng khi đo lường hiệu suất của một mô hình.

		Classifier Prediction	
		Positive	Negative
Actual Value	Positive	True Positive	False Negative
	Negative	False Positive	True Negative

Hình 2: Ma trận nhầm lẫn mẫu

1. Dương tính thực: Cả giá trị dự đoán và đầu ra thực tế đều là 1. 2. Âm tính đúng: Cả giá trị dự đoán và đầu ra thực tế đều bằng 0. 3. Dương tính giả: Giá trị dự đoán là 1, nhưng đầu ra thực tế là 0. Bốn. Phủ định sai: Giá trị dự đoán là 0, nhưng đầu ra thực tế là 1. Ma trận nhầm lẫn cơ bản để phân loại nhị phân. Các đường chéo của ma trận phải luôn chứa giá trị lớn nhất, vì các đường chéo đại diện cho một mẫu được phân loại tốt. Trong phân loại nhiều lớp, mỗi lớp đại diện cho một hàng và cột ma trận.

C. Độ chính xác của mô hình và đường cong suy giảm mô hình:

Những đường cong này, còn được gọi là đường cong học tập, được sử dụng phổ biến nhất cho các mô hình học từng bước theo thời gian, chẳng hạn như mạng lưới thần kinh. Các mạng thần kinh này đại diện cho việc đào tạo và đánh giá dữ liệu xác thực và bạn có thể thấy mô hình có thể được đào tạo và khái quát hóa tốt như thế nào. Đường cong tổn thất mô hình biểu thị điểm tối thiểu hóa (tổn thất). Nói cách khác, điểm số càng thấp thì hiệu suất của mô hình càng tốt. Đường cong chính xác được sử dụng trong học máy biểu thị điểm số tối đa. Trong những trường hợp này, chúng ta có thể nói rằng điểm số càng cao thì hiệu suất của mô hình càng tốt. Một đường cong tổn thất mô hình tốt là một trong đó các tổn thất huấn luyện và xác thực được giảm xuống để đạt đến điểm ổn định và có khoảng cách tối thiểu giữa các giá trị tổn thất cuối cùng. Đường cong chính xác và ổn định hơn trong đào tạo và xác thực, với khoảng cách tối thiểu giữa giá trị độ chính xác cuối cùng là mô hình tốt nhất.

TRONG. HỀ ĐỀ XUẤT

Máy theo dõi yoga thể dục dựa trên AI thường được thiết kế để sử dụng thông qua các thiết bị có camera có thể ghi tọa độ góc và chụp nhiều hình ảnh hơn trong quá trình thực hiện bài tập. Thuật toán thông thường cho trình theo dõi dựa trên ước tính tư thế của con người là: Khi người dùng bắt đầu sử dụng trình theo dõi yoga thể dục, máy ảnh sẽ ghi lại chuyển động của họ trong khi tập luyện. Ảnh đã chụp và tọa độ góc được chia thành các khung riêng lẻ và được xử lý bằng mô hình ước tính tư thế con người. Mô hình này phát hiện các điểm chính trong cơ thể người dùng và tạo thành một "bộ xương" ảo ở kích thước 2D hoặc 3D. Khung ảo được phân tích bằng các quy tắc dựa trên hình học hoặc các phương tiện khác để xác định các lỗi trong phương pháp làm bài tập (nếu có). Người dùng nhận được mô tả về các lỗi đã mắc phải và các đề xuất để loại bỏ chúng.

Hệ thống được đề xuất được triển khai trong python bằng thư viện OpenCV và chạy trên CPU Lenovo Intel Core i3, RAM 4 GB và hệ điều hành Windows 10 64-bit. Một bộ dữ liệu chứa 84 tư thế yoga asana điển hình được hệ thống chọn bằng cách sử dụng webcam thông thường và được cung cấp công khai. Một cách tiếp cận lai mới dựa trên phân loại học máy và học sâu. Bước một bao gồm, một máy vectơ hỗ trợ (SVM) được lên kế hoạch. Trình phân loại này sử dụng khả năng dự đoán của máy học để cải thiện hiệu suất của các thuật toán ML. Bước thứ hai là một mạng lưới thần kinh tích chập, nắm bắt bộ xương của các tư thế và tư thế mục tiêu của người dùng, đồng thời so sánh hai tư thế để có được sự tương đồng. Tôi đã nhập các thư viện như Hệ điều hành, Thời gian, Bàn phím, Mạng và Mediapipe. Giả thuyết của chúng tôi là tọa độ của các bộ phận cơ thể khác nhau trên cơ thể người từ các hình ảnh chứa thông tin để xác định xem tư thế có được thực hiện đúng hay không.

VI. BÁO CÁO VẤN ĐỀ

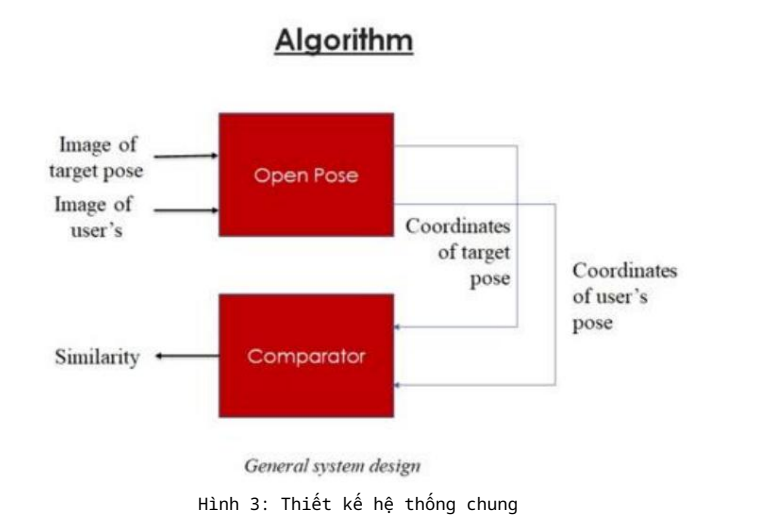
Gần đây, tình cờ bắt gặp và chọn ra mức độ phổ biến của chuyển động con người rất khó khăn và với nhịp sống nhanh của chúng ta trong thời gian gần đây, con người thường chọn tập luyện tại nhà, tuy nhiên, họ cảm thấy muốn có một người hướng dẫn kiểm tra hình thức tập luyện của mình. Vì những nguồn đó không phải lúc nào cũng có vấn đề, nên sự nổi tiếng của con người có thể được sử dụng để tạo ra một phần mềm tập luyện theo nhịp độ cá nhân cho phép con người nghiên cứu, tập thể dục và tập luyện đúng cách. Thiết bị được đề xuất được sử dụng với Mạng thần kinh chuyển đổi để nhận biết hành động của con người dựa trên phân loại tư thế yoga bằng cách sử dụng xử lý hình ảnh và

A. Mục tiêu

Mục tiêu của ước tính tư thế là để theo dõi chuyển động của các đối tượng con người cho các bài tập riêng biệt. Để xem xét việc sử dụng một mô hình kỹ năng trước, với mục đích sử dụng đánh giá hình học để nhận thức được bài tập đạt được trong video, có thể là trong thời gian thực. Để xem xét thành phần cung cấp khả năng vectơ diện tích để tính góc giữa các phần khung. Để đánh giá cơ chế sinh học của con người trong một video tập thể dục hoặc học tập và đưa ra nhận xét về thói quen của họ.

B. Thuật toán

Một phương pháp lai mới hoàn toàn dựa trên kiến thức thu được của hệ thống về các bộ phận loại đã được đề xuất. Đầu tiên là hệ thống vectơ hướng dẫn (SVM), sử dụng hệ thống thu được kiến thức về dự đoán (ML) để nâng cao hiệu suất tổng thể của các thuật toán ML. Kỹ thuật 2d là áp dụng mạng nơ-ron tích chập (CNN) để lấy bộ xương người của người tiêu dùng và tư thế mục tiêu rồi kiểm tra chúng để đưa ra điểm tương đồng.

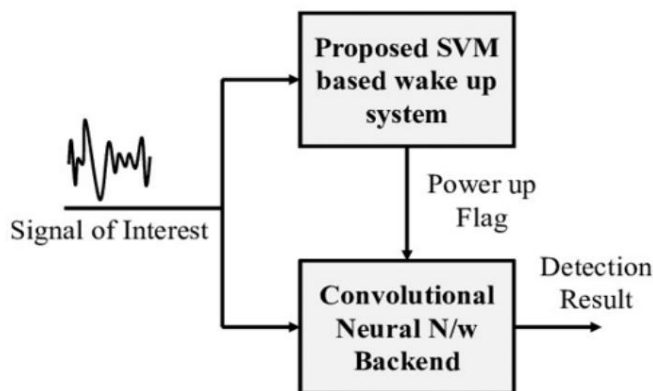


Hình 3: Thiết kế hệ thống chung



## C. Máy véc tơ hỗ trợ (SVM):

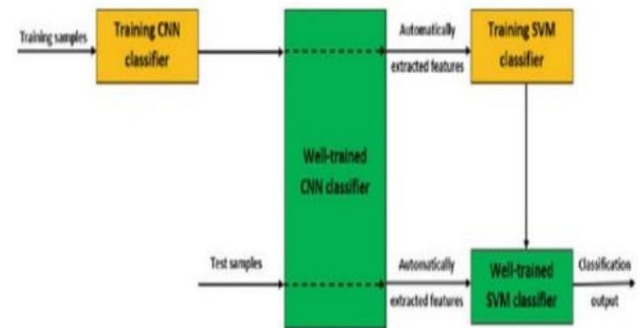
SVM là một mô hình học công cụ được giám sát, vốn dĩ là một bộ phân loại -elegance. Tuy nhiên, vì hầu hết các vấn đề bao gồm nhiều huấn luyện, nên một SVM đa lớp thường được sử dụng. Tài liệu SVM đa lớp có nhiều bộ phân loại sang trọng và phân biệt các bộ phân loại dựa trên ý tưởng về nhân tốt so với phần còn lại (một đối với phần còn lại hoặc một đối với tất cả) hoặc giữa từng cặp đào tạo (một đối với một). SVM thực hiện lớp thông qua việc phát triển một siêu phẳng theo cách sao cho sự tách biệt giữa đào tạo càng nhiều càng tốt. Một SVM mặc định đã có kiến thức về số liệu thống kê đào tạo với hạt nhân hàm cơ sở radial (rbf). Rbf là hạt nhân mặc định và nổi tiếng nhất có thể là hàm cơ sở xuyên tâm gaussian. Nó cung cấp tính linh hoạt hơn so với các hạt nhân đặc biệt, tuyến tính và đa thức. Tỷ lệ của tham số lề để dành C là 1 và chức năng lựa chọn là một đối với phần còn lại. Các điểm chính được chụp bằng OpenPose được sử dụng làm tài năng cho SVM. 18 điểm chính này được thể hiện thông qua tọa độ X và Y, tạo nên loại tài năng mở rộng chung là 36 ( $18 * 2$ ). Số liệu thống kê được định hình lại để làm cho các loại mẫu mở rộng bằng nhau.



Hình 4: Kiến trúc SVM

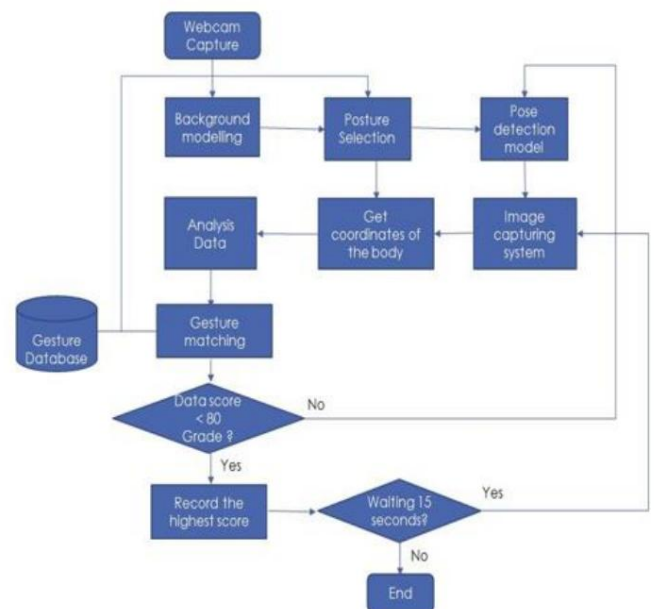
## D. Mạng thần kinh tích chập (CNN):

Cộng đồng thần kinh tích chập (CNN) là một phân đoạn thu thập kiến thức chuyên sâu về cộng đồng. CNN đại diện cho sự phát triển cơ bản trong nhận dạng hình ảnh. Nó hoạt động như Đầu vào - Chúng tôi có chế độ chụp ảnh trực tiếp của người dùng. Convolution - Trong bước này, chúng ta có thể vẽ các bộ dò chức năng, hoạt động như các bộ lọc cộng đồng thần kinh. Sau đó vẽ tranh trên sơ đồ chức năng. Tổng hợp - Trong lớp này, chúng tôi thu nhỏ ngăn xếp ảnh thành một kích thước nhỏ hơn. Pooling kết thúc sau khi đặt nó qua lớp phát hành. Làm phẳng - Chúng tôi chuyển đổi dữ liệu thành mảng 1 chiều để chèn nó vào lớp tiếp theo. Kết nối đầy đủ - Chúng tôi đã hoàn thành giáo dục cộng đồng và có thể bắt đầu mong đợi và kiểm tra hiệu suất tổng thể của phân vùng.



Hình 5: Kiến trúc CNN.

## VII. KIẾN TRÚC HỆ THỐNG



Hình 6: Kiến trúc hệ thống

Tư thế - Nhìn chung, PoseNet sẽ di chuyển trả về một đối tượng tư thế có chứa danh sách các yếu tố chính và điểm tự bảo hành văn bằng đối tượng cho một người được phát hiện. Điểm tự đảm bảo của tư thế - Xác định mức tự đảm bảo tổng thể trong phạm vi ước tính của một tư thế, nằm trong khoảng từ 0,0 đến 1,0. Nó có thể được sử dụng để che những tư thế được coi là không đủ mạnh. Keypoint - Một phần tư thế của một người trong đó khía cạnh quan trọng của cơ thể được dự đoán, bao gồm mũi, tai phải, đầu gối trái, bàn chân phải, v.v. Điểm Tin cậy Keypoint - Xác định khả năng tự bảo đảm của tính năng keypoint được dự đoán là chính xác, khác nhau giữa 0,0 và 1,0. Nó có thể được sử dụng để che đậy các yếu tố chính được coi là không đủ mạnh. Vị trí điểm chính - Tọa độ 2D x và y bên trong hình ảnh đầu vào phù hợp trong đó khía cạnh chính đã được phát hiện.

VIII. KHẢO SÁT VĂN HỌC

[1] “Phương pháp đánh giá tư thế Yoga bằng cách sử dụng tính năng phát hiện tư thế để tự học” của MC Thar, KZN Bài viết này đề xuất một phương pháp đánh giá tư thế Yoga đối với việc sử dụng tính năng phát hiện tư thế để hỗ trợ việc tự làm chủ Yoga. Bài báo này đã đề xuất một Hệ thống đánh giá hiệu suất như Hệ thống đào tạo tư thế Yoga để hỗ trợ việc tự làm chủ Yoga. Bài viết này đưa ra cách khám phá các tư thế yoga và việc sử dụng khám phá tư thế để hỗ trợ việc tự kiểm tra yoga.

[2] “Yoga thời gian thực đánh giá cao việc sử dụng kỹ năng làm chủ sâu” Bởi SK Yadav, A. Singh. Một phiên bản làm chủ kết hợp cường độ cao đã đề xuất sử dụng CNN và LSTM để hiển thị yoga trong các bộ phim thời gian thực, trong đó lớp CNN được sử dụng để trích xuất các khả năng từ đối tượng quan trọng của mọi người được xác định ở tư thế mở và quan sát thông qua LSTM để cung cấp tạm thời phỏng đoán. Bài báo này chỉ ra một ứng dụng yoga trợ lý di động hoàn toàn dựa trên các kiểu thu nhận khóa của con người để trò chuyện video.

[3] “Ứng dụng Yoga dành cho thiết bị di động dựa trên tính năng phát hiện yoga” của Sylvie. Các tác giả cho thấy ứng dụng di động của trợ lý yoga dựa trên mô hình cá nhân nơi người hướng dẫn hướng dẫn và giám sát học viên của họ tập yoga bằng trò chuyện video. Một mô hình học tập chuyên sâu đã được đề xuất sử dụng CNN và LSTM để nhận dạng yoga trong các video thời gian thực.

[4] “Tư thế Yoga học ML trong chuỗi video” của Jozsef. Cách tiếp cận vấn đề là nghiên cứu mô hình sản xuất các mẫu chuyển động bình thường bằng cách sử dụng nhiều nguồn với khả năng kiểm soát hạn chế. Với hiệu suất không được mã hóa, chúng tôi đề xuất hai bộ mã hóa tự động để chúng có thể hoạt động mà không cần hướng dẫn tối thiểu.

[5] “Phát hiện thời gian thực trong khu vực đông đúc” của Ammar Ladjaillia. Một phương pháp mã hóa động hoàn toàn không được giám sát để phát hiện các sự kiện bất thường trong video dựa trên khả năng xây dựng trực tuyến các tín hiệu truy vấn từ các từ điển sự kiện đã học.

IX. THỰC HIỆN MÔ HÌNH VÀ KẾT QUẢ

A. Hệ thống vectơ hỗ trợ (SVM):

SVM là một tiện ích được giám sát để biết phiên bản này vốn dĩ là một bộ phân loại -elegance. Tuy nhiên, đối với các vấn đề tối đa có trong quá trình đào tạo, một SVM đa lớp thường được sử dụng. Một tài liệu SVM đa lớp chứa một số bộ phân loại thanh lịch và phân biệt các bộ phân loại dựa trên ý tưởng về nhãn tuyệt vời. Thư giãn (thư giãn một chơi một hoặc một chơi tất cả) hoặc giữa các cặp luyện tập (một chơi một). SVM chơi loại với sự trợ giúp của việc sử dụng phát triển một siêu phẳng theo bất kỳ cách nào sao cho sự tách biệt giữa các bài tập càng lớn càng tốt.

Một SVM mặc định đã được giáo dục tại các bản ghi giáo dục với hạt nhân đặc tính cơ sở hướng tâm (RBF). Rbf là hạt nhân nổi tiếng mặc định và tối đa là đặc tính của nền tảng xuyên tâm Gaussian. Nó cung cấp thêm tính linh hoạt so với các hạt nhân khác, tuyến tính và đa thức. Các điểm chính được nắm bắt trong quá trình sử dụng OpenPose được sử dụng làm khả năng cho SVM. 18 điểm chính này được thể hiện với sự trợ giúp của việc sử dụng tọa độ X và Y, tạo ra đầy đủ các khả năng

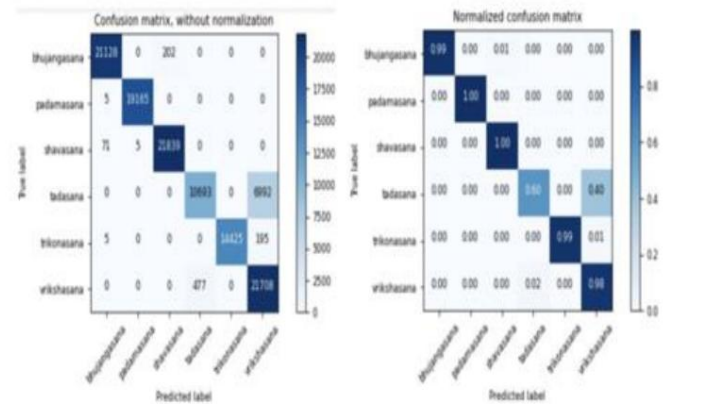
36(18\*2). Các bản ghi được định hình lại để làm cho nhiều loại mẫu bằng nhau.

Kết quả:

- Đào tạo chính xác: 0,9953
- Độ chính xác xác thực: 0,9762
- Độ chính xác kiểm tra: 0,9319

Phân tích:

Độ chính xác huấn luyện của mô hình khá cao ở mức 0,99. Có một sự sụt giảm nhỏ trong quá trình xác thực và kiểm tra độ chính xác, tuy nhiên, kết quả vẫn tốt. Chúng ta sẽ thấy trong ma trận nhầm lẫn rằng hầu hết tất cả các loại được phân loại chính xác ngoài tadasana (nguyên nhân núi). Trong số 17.685 khung cho tadasana, 6992 được phân loại sai thành trikonasana (tư thế cái cây) và tương tự, có một số phân loại sai cho vrikshasana. Điều này có thể là do sự giống nhau trong các tư thế vì mỗi người trong số họ cần một vị trí đứng và đồng thời, sự hình thành tư thế ban đầu cũng tương tự nhau.



Hình 7: SVM

B. Mạng thần kinh tích chập (CNN):

Biểu mẫu đầu vào là một cặp biểu thị 32 điểm chính có tọa độ X và Y. Chuẩn hóa hàng loạt được áp dụng cho đầu ra của lớp CNN để mô hình hội tụ nhanh hơn. Chúng tôi thậm chí có xu hướng có một lớp bỏ học ngăn chặn quá mức phù hợp bằng cách giảm một cách bữa bãi một số trọng số. Hiệu suất được sử dụng là thước đo tuyến tính đã hiệu chỉnh được áp dụng để trích xuất tính năng trên các điểm chính của mọi khung hình. Đầu ra cuối cùng là phẳng trước khi được chuyển đến lớp dày đặc với các đơn vị Liên kết kích hoạt tối đa mềm ở bất kỳ đâu mỗi đơn vị biểu thị cơ hội yoga được tạo theo thuật ngữ entropy chéo cho tất cả các danh mục.

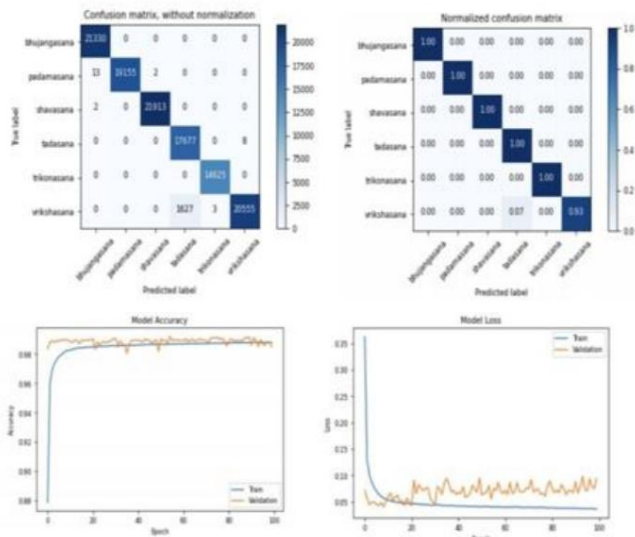
Điều này có thể được sử dụng vì nó cho phép đo hiệu suất của đầu ra của lớp kết nối dày đặc với kích hoạt soft-max. Hàm mất mát này được sử dụng để phân loại nhiều loại và vì chúng tôi có nhiều lớp tư thế yoga, nên sử dụng entropy chéo phân loại là điều hợp lý. Cuối cùng, để quản lý tốc độ học, một trình tối ưu hóa có tốc độ học ban đầu là 0,0001 được sử dụng. Số lượng tổng thể của các kỷ nguyên mà mô hình được đào tạo là 100.

Kết quả:

- Đào tạo chính xác: 0,9841
- Độ chính xác thực: 0,9910
- Độ chính xác kiểm tra: 0,9868

Phân tích:

Việc huấn luyện, xác nhận và kiểm tra độ chính xác của mô hình gần như giống nhau, chỉ khoảng 0,99. Ma trận nhầm lẫn bổ sung cho thấy mô hình sẽ phân loại chính xác tất cả các mẫu, ngoại trừ một số mẫu trong trikonasana bị phân loại sai thành tadasana, dẫn đến 93 lần phân loại sai trên CNN. Tuy nhiên, đường cong tổn thất mô hình cao hơn cho thấy sự gia tăng trong tổn thất xác nhận và giảm tổn thất đào tạo, điều này cho thấy có một số trang bị thừa.



Hình 8: CNN

x. KẾT QUẢ VÀ PHÂN TÍCH

Độ chính xác của mô hình huấn luyện, xác nhận và kiểm tra gần như giống nhau, gần bằng 0,99. Ma trận nhầm lẫn cho thấy rằng mô hình sẽ phân loại chính xác tất cả các mẫu, ngoại trừ một số mẫu trong trikonasana bị phân loại sai thành tadasana, dẫn đến độ chính xác chín mươi ba cho vrikshasana. Có ít phân loại sai hơn trong CNN. Tuy nhiên, đường cong tổn thất mô hình ở trên cùng cho thấy sự gia tăng tổn thất xác thực và giảm tổn thất huấn luyện, điều này cho thấy có một số trang bị thừa.

Độ chính xác huấn luyện của mô hình cao ở mức 0,98. Có một mất mát nhỏ trong quá trình xác nhận và kiểm tra độ chính xác, tuy nhiên, kết quả vẫn tốt. Chúng ta có thể thấy trong ma trận nhầm lẫn rằng hầu hết tất cả các danh mục đều được phân loại chính xác ngoài tadasana (tạo núi). Trong số 17.695 khung hình cho tadasana, 6982 được phân loại sai thành trikonasana và tương tự, có một số phân loại sai cho vrikshasana. Điều này có thể là do sự giống nhau trong các tư thế vì mỗi người trong số họ cần một vị trí đứng và thêm vào đó, sự hình thành tư thế ban đầu là tương tự nhau.

A. Trang chủ:

Mô tả: Khi truy cập trang chủ, lời chào bằng lời nói được cung cấp cho người dùng. Người dùng có thể nhấp vào nút BẮT ĐẦU để tiếp tục phần thực hành, điều hướng đến trang THỰC HÀNH NGAY BÂY GIỜ.

B. Trang Thực Hành Ngay:

Mô tả: Trang thực hành ngay bây giờ sẽ hiển thị tất cả các tư thế do hệ thống cung cấp và các nguyên tắc mà người dùng phải tuân theo. Tại đây, người dùng có thể chọn một tư thế cụ thể hoặc có thể nhấp vào nút BẮT ĐẦU để thực hiện tất cả các tư thế yoga theo trình tự.

C. Cố định Trang Vị trí:

Mô tả: hiển thị trang cố định vị trí trong đó hướng dẫn người dùng căn chỉnh các điểm chính của mắt và mắt cá chân, do đó vị trí của người dùng được cố định trước máy ảnh. Nó sẽ được thực hiện để đảm bảo toàn bộ cơ thể của người dùng (từ đầu đến chân) có thể nhìn thấy trong máy ảnh.

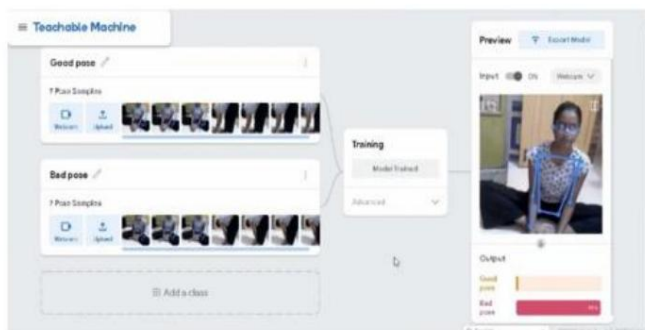
D. Thực hành một tư thế

Mô tả: Trong khi thực hành một tư thế, trước tiên, tư thế của người dùng được xác định và các điểm chính được tạo cũng như bộ xương được vẽ trên canvas video. Hướng dẫn được cung cấp cho người dùng để thực hiện tư thế yoga một cách chính xác. Sau khi thực hiện thành công tư thế yoga, hệ thống sẽ yêu cầu người dùng giữ nguyên tư thế và hít thở sâu.

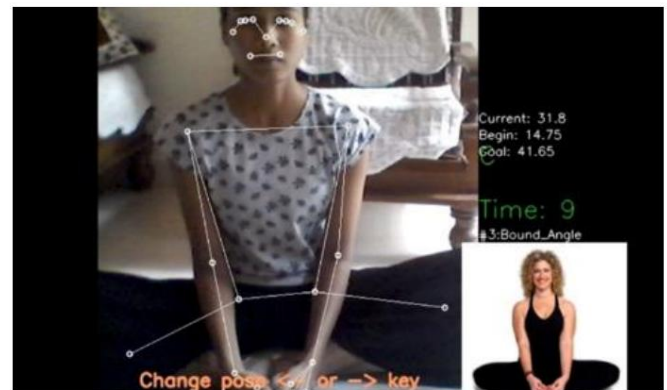
XI. PHẠM VI TƯƠNG LAI

Mô hình được đề xuất hiện phân loại 6 asana yoga hiệu quả nhất. Có một số asana yoga, và do đó, việc phát triển một phiên bản đánh giá tư thế có thể thành công cho tất cả các asana là điều khó khăn. Tập dữ liệu có thể được cải thiện bằng cách bổ sung thêm các tư thế yoga được thực hiện bởi những người hiện không còn hiệu quả nhất ở vị trí trong nhà mà còn ở ngoài trời. Hiệu suất tổng thể của các mô hình phụ thuộc vào tính năng ước lượng tư thế OpenPose đặc biệt có thể không hoạt động tốt trong các trường hợp chồng chéo giữa các người hoặc chồng chéo giữa các bộ phận khung. Một công cụ di động để tự học và dự đoán thời gian thực có thể được thực hiện cho hệ thống này. Bức tranh này thể hiện danh tiếng của trò tiêu khiển đối với các ứng dụng thực tế. Một phương pháp tương tự như thế này có thể được áp dụng cho danh tiếng của tư thế trong các trách nhiệm bao gồm thể thao, giám sát, chăm sóc sức khỏe, v.v. Bản thân việc ước tính tư thế đa ký tự là một vấn đề hoàn toàn mới và có nhiều phạm vi nghiên cứu. Có rất nhiều tình huống mà việc ước tính tư thế của một nhân vật giờ đây có thể không còn đủ nữa, chẳng hạn như việc ước tính tư thế trong các tình huống đồng đức có thể có nhiều hơn một người như một cách để kiểm soát và tìm ra tư thế của mỗi cá nhân. Ngoài ra, rất nhiều thứ bao gồm nền, ánh sáng, các hình chồng lên nhau, v.v. đã được đề cập trước trong cuộc khảo sát này có thể khiến việc ước lượng tư thế đa nhân vật trở nên khó khăn.

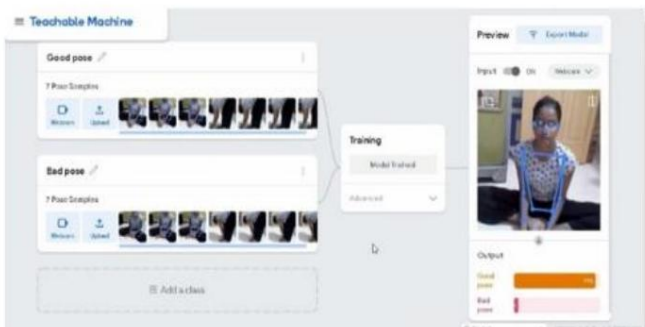




Hình 9: Tư thế Yoga 1



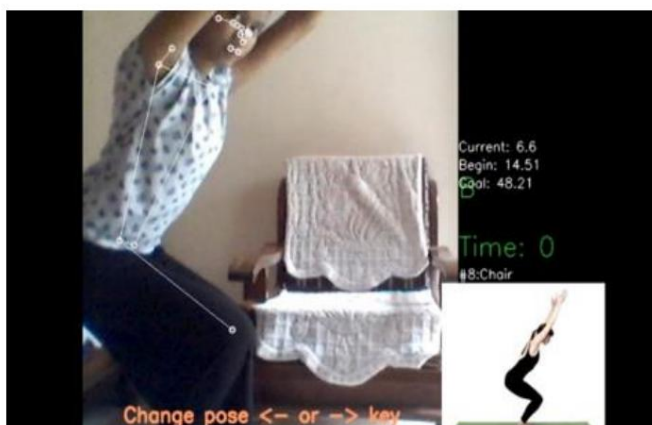
Hình 9: Tư thế Yoga 5



Hình 9: Tư thế Yoga 2



Hình 9: Tư thế Yoga 3



Hình 9: Tư thế Yoga 4

## XII. PHẦN KẾT LUẬN

Lớp CNN phân phối thời gian phát hiện các mẫu giữa các điểm chính trong một khung và cả đánh giá SVM trong các khung mới nhất để ghi nhớ các khung trước đó, kết quả là hệ thống thậm chí còn mạnh hơn rất nhiều bằng cách giảm thiểu lỗi do phát hiện điểm chính sai. Bởi vì các khung ảnh Yoga là tuần tự. Chúng tôi có xu hướng lên kế hoạch cho một hệ thống nhận dạng Yoga bằng camera RGB thông thường. Tập dữ liệu được thu thập bằng camera kỹ thuật số HD 1080p của Logitech dành cho 15 người (mười nam và 5 nữ) và được tạo ra công khai. Khung dựa trên máy học loại bỏ các tùy chọn bổ sung các asana mới nhất bằng cách chỉ cần chuẩn bị mô hình với dữ liệu mới.

Chúng tôi có xu hướng áp dụng lớp CNN phân phối theo thời gian để phát hiện các mẫu giữa các điểm chính trong một khung hình và LSTM để nghiên cứu các mẫu được tìm thấy trong các khung hình gần đây. Sử dụng LSTM cho bộ nhớ của các khung hình trước đó và bỏ phiếu để khử nhiễu, kết quả giúp hệ thống trở nên mạnh mẽ hơn bằng cách giảm thiểu lỗi do phát hiện điểm chính sai. Vì các khung hình của video Yoga là tuần tự. Cách tiếp cận tương tự có thể được sử dụng để nhận dạng tư thế trong các tác vụ khác nhau như thể thao, chăm sóc sức khỏe và phân loại hình ảnh.

## NGƯỜI GIỚI THIỆU

- [1]. L.Sigal. "Ước tính tư thế con người", Ency. của điện toán. Tầm nhìn, Springer 2011.
- [2]. S. Yadav, A. Singh, A. Gupta và J. Raheja, "Nhận dạng yoga theo thời gian thực bằng cách sử dụng học sâu", Neural Comput. và Appl., tháng 5 năm 2019. [Trực tuyến]. Có sẵn: <https://doi.org/10.1007/s00521-019-04232-7>
- [3]. U. Rafi, B. Leibe, J.Gall, và I. Kostrikov, "Một mạng tích chập hiệu quả để ước lượng tư thế con người", British Mach. Vision Conf S. Haque, A. Rabby, M. Laboni, ., 2016.
- [4]. N. Neehal, và S. Hossain, "ExNET: mạng thần kinh sâu để phát hiện tư thế tập thể dục", Xu hướng gần đây trong quá trình hình ảnh. và Patter Recog., 2019.
- [5]. M. Islam, H. Mahmud, F. Ashraf, I. Hossain và M. Hasan, "Nhận dạng tư thế yoga bằng cách phát hiện các điểm khớp của con người trong thời gian thực bằng cách sử dụng microsoft kinect", IEEE Khu vực 10 Humanit. Công nghệ. Conf., trang 668-67, 2017.

- [6]. S. Patil, A. Pawar, và A. Peshave, "Người hướng dẫn yoga: trực quan hóa và phân tích bằng thuật toán SURF", Proc. Hệ thống điều khiển IEEE. Hội thảo nghiên cứu sau đại học, trang. 43-46, 2011.
- [7]. W. Gong, X. Zhang, J. Gonzalez, A. Sobral, T. Bouwmans, C. Tu, và H. Zahzah, "Ước tính tư thế con người từ hình ảnh bằng một mắt: một cuộc khảo sát toàn diện", Sensors, Basel, Thụy Sĩ, tập. 16, 2016.
- [số 8]. G. Ning, P. Liu, X. Fan và C. Zhan, "Cách tiếp cận từ trên xuống để ước tính và theo dõi tư thế con người rõ ràng", Hội thảo ECCV, 2018.
- [9]. A. Gupta, T. Chen, F. Chen và D. Kimber, "Các hệ thống và phương pháp lượng tử thể cơ thể con người", bằng sáng chế Hoa Kỳ, 7,925,081 B2, 2011.
- [10]. H. Sidenbladh, M. Black và D. Fleet, "Theo dõi ngẫu nhiên hình người 3D bằng cách sử dụng chuyển động hình ảnh 2D", Proc 6th European Conf. Thị giác máy tính, 2000.
- [11]. A. Agarwal và B. Triggs, "Tư thế người 3D từ bóng bằng hồi quy véc tơ liên quan", Hội nghị Quốc tế. trên Mẫu thị giác máy tính
- [12]. M. Li, Z. Zhou, J. Li và X. Liu, "Ước tính tư thế từ dưới lên của nhiều người với ràng buộc hộp giới hạn", 24th Intl. Conf. Nhận dạng mẫu, 2018.
- [13]. Z. Cao, T. Simon, S. Wei và Y. Sheikh, "OpenPose: ước tính tư thế 2D nhiều người trong thời gian thực bằng cách sử dụng các trường sở thích bộ phận", Proc. Hội nghị IEEE lần thứ 30 Thị giác Máy tính và Nhận dạng Mẫu, 2017.
- [14]. A. Kendall, M. Grimes, R. Cipolla, "PoseNet: mạng tích chập dành cho việc tái định vị máy ảnh 6-DOF thời gian thực", IEEE Intl. Conf. Thị giác máy tính, 2015.
- [15]. S. Kreiss, L. Bertoni, và A. Alahi, "PifPaf: các trường tổng hợp để ước tính tư thế con người", IEEE Conf. Thị giác máy tính và Nhận dạng mẫu, 2019.
- [16]. P. Dar, "AI guardman - một ứng dụng học máy sử dụng ước tính tư thế để phát hiện những kẻ trộm cắp". [Trực tuyến]. Có sẵn 2018
- [17]. D. Mehta, O. Sotnychenko, F. Mueller và W. Xu, "XNect: ước tính tư thế con người 3D nhiều người trong thời gian thực bằng một camera RGB duy nhất", ECCV, 2019.
- [18]. A. Lai, B. Reddy và B. Vlijmen, "Yog.ai: deep learning [Trực tuyến]. yoga". Vì Có sẵn: <http://cs230.stanford.edu/projects/winter2019>
- [19]. M. Dantone, J. Gall, C. Leistner, "Ước tính tư thế con người bằng cách sử dụng các biến hồi quy khớp phụ thuộc vào các bộ phận cơ thể", Proc. Hội nghị IEEE Nhận dạng mẫu thị giác máy tính, 2013.
- [20]. A. Mohanty, A. Ahmed, T. Goswami, "Nhận dạng tư thế mạnh mẽ bằng cách sử dụng học sâu", Adv. trong Hệ thống thông minh. và Comput, Singapore, trang 93-105, 2017.
- [21]. P. Szczuko, "Mạng lưới thần kinh sâu để ước tính tư thế con người từ hình ảnh độ sâu có độ phân giải rất thấp", Công cụ đa phương tiện và Ứng dụng, 2019.
- [22]. M. Chen, M. Low, "Ước tính tư thế con người định kỳ", [Trực tuyến]. Có sẵn: <https://web.stanford.edu/class/cs231a/prevprojects201>
- 6/ trận chung kết
- [23]. K. Pothanackar, "Nhận dạng hành động của con người bằng cách sử dụng CNN và LSTMNN với mô hình chú ý", Tạp chí Công nghệ Sáng tạo Quốc tế. và Khám phá tiếng Anh, 2019.
- [24]. N. Nordsborg, H. Espinosa, "Ước tính mức tiêu hao năng lượng trong quá trình bơi trườn về phía trước bằng phương pháp đo gia tốc", Procedia Eng., 2014.
- [25]. P. Pai, L. Changliao, K. Lin, "Phân tích trò chơi bóng rổ bằng máy vectơ hỗ trợ với mô hình cây quyết định", Neural Comput. Ứng dụng, 2017. [26] S. Patil, A. Pawar, A. Peshave, "Gia sư yoga: hình dung và phân tích bằng thuật toán SURF", Proc. Hệ thống điều khiển IEEE. tốt nghiệp. Nghiên cứu Colloquium, 2011.
- [26]. W. Wu, W. Yin, F. Guo, "Hệ thống chuyên gia học tập và tự hướng dẫn cho yoga", Proc. quốc tế Hệ thống làm việc thông minh Ứng dụng, 2010.
- [27]. E. Trejo, P. Yuan, "Nhận biết các tư thế yoga thông qua hệ thống tương tác với thiết bị kinect", Quốc tế. Conf. Khoa học người máy và tự động hóa, 2018.