





## Bài báo nghiên cứu

# Ước tính tư thế Yoga và tạo phản hồi bằng cách sử dụng Học kỹ càng

Vivek Anand Thoutam,<sup>1</sup> Anugrah Srivastava,<sup>1</sup> Tapas Badal,<sup>1</sup> Vipul Kumar Mishra,<sup>1</sup> GR Sinha <sup>2</sup>, Aditi râu <sup>3</sup>, Harshit Bhardwaj <sup>3</sup> và Manish Raj <sup>4</sup>

<sup>1</sup>Khoa Kỹ thuật Khoa học Máy tính, Đại học Bennett, Greater Noida, Ấn Độ

<sup>2</sup>MIIT, Mandalay, Myanmar

<sup>3</sup>Trường Khoa học và Kỹ thuật Máy tính, Đại học Galgotias, Greater Noida, Ấn Độ

<sup>4</sup>Khoa Kỹ thuật Khoa học Máy tính, Đại học Bennett, Greater Noida, Ấn Độ

Thư từ nên được gửi tới GR Sinha; gr\_sinha@miit.edu.mm

Nhận ngày 12 tháng 12 năm 2021; Sửa đổi ngày 17 tháng 2 năm 2022; Chấp nhận ngày 26 tháng 2 năm 2022; Xuất bản ngày 24 tháng 3 năm 2022

Biên tập học thuật: Vijay Kumar

Bản quyền © 2022 Vivek Anand outam et al. là một bài báo truy cập mở được phân phối theo Giấy phép ghi nhận tác giả Creative Commons, cho phép sử dụng, phân phối và sao chép không hạn chế trong bất kỳ phương tiện nào, miễn là tác phẩm gốc được trích dẫn chính xác.

Yoga là một môn tập luyện 5000 năm tuổi được phát triển ở Ấn Độ cổ đại bởi nền văn minh Indus-Sarasvati. Từ yoga có nghĩa là sự liên kết sâu sắc và sự hợp nhất của tâm trí với cơ thể. Nó được sử dụng để giữ cho cả tâm trí và cơ thể cân bằng trong tất cả các bậc ngót của cuộc sống bằng asana, thiền định và một số kỹ thuật khác. Ngày nay, yoga đã thu hút được sự chú ý trên toàn thế giới do mức độ căng thẳng gia tăng trong lối sống hiện đại, và có rất nhiều phương pháp hoặc tài nguyên để học yoga. Yoga có thể được thực hành tại các trung tâm yoga, thông qua các gia sư cá nhân và cũng có thể tự học với sự trợ giúp của Internet, sách báo, clip ghi lại, v.v. có thể không có sẵn mọi lúc.

Nhưng khi tự học, không thể tìm thấy một tư thế sai. Tư thế không đúng có thể gây hại cho sức khỏe của một người, dẫn đến đau cấp tính và các vấn đề mãn tính lâu dài. Trong bài báo này, các kỹ thuật dựa trên học sâu được phát triển để phát hiện tư thế yoga không đúng. Với phương pháp này, người dùng có thể chọn tư thế luyện tập mong muốn và có thể tải lên các video đã ghi về tư thế luyện tập yoga của họ. Tư thế của người dùng được gửi đến các mô hình huấn luyện tạo ra các góc bắt đầu được phát hiện giữa tư thế thực tế và tư thế của người dùng. Với những kết quả đầu ra này, hệ thống sẽ khuyên người dùng cải thiện tư thế bằng cách chỉ định vị trí sai của tư thế yoga. Phương pháp đề xuất được so sánh với một số phương pháp tiên tiến nhất và nó đạt được độ chính xác vượt trội là 0,9958 trong khi yêu cầu độ phức tạp tính toán ít hơn.

## 1. Giới thiệu

Giống như mọi bài tập, điều quan trọng nhất là thực hành các tư thế yoga một cách chính xác vì bất kỳ tư thế bắt đầu nào cũng không hiệu quả và có xu hướng gây hại. gợi ý nên có một người hướng dẫn xung quanh khi thực hiện yoga. Không phải lúc nào bạn cũng có thể có một người hướng dẫn hoặc tham gia các lớp học yoga với lối sống ngày nay. Một hệ thống dựa trên AI giúp xác định các tư thế yoga và đưa ra phản hồi hoặc gợi ý cho người dùng. Các hướng dẫn này giúp người dùng cải thiện các tư thế của họ sao cho hiệu quả và không gây bất lợi. Các thách thức trong dự án này là các điểm chính phải được phát hiện mà không có bất kỳ điểm nào bị thiếu và các mô hình phải hoạt động bình thường ngay cả khi các bộ phận cơ thể bị biến dạng.

chồng lên nhau. Các đề xuất nên được đưa ra một cách chính xác vì những thay đổi nhỏ có thể gây ra kết quả có hại. Các tư thế trong bộ dữ liệu được sử dụng cho dự án này nên được thực hiện bởi các chuyên gia. Người mẫu nên phân loại chính xác các tư thế, ngay cả khi chúng gần giống nhau với một số khác biệt nhỏ.

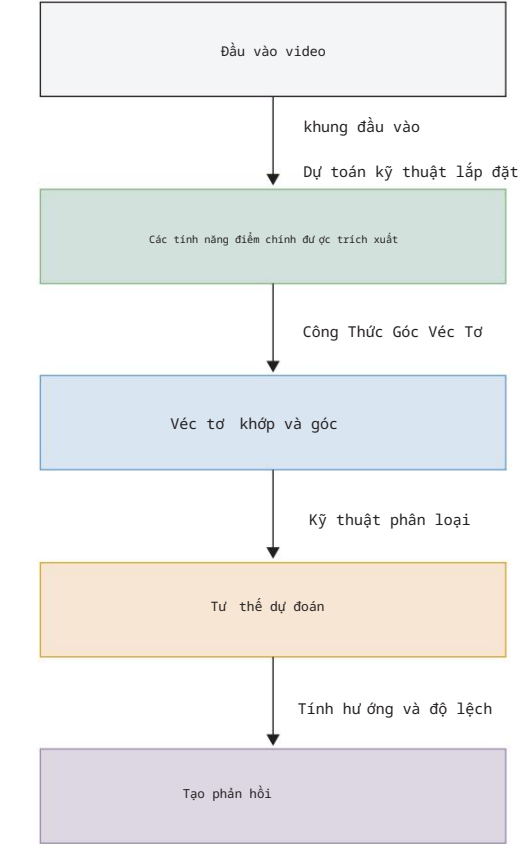
Các phương pháp tự huấn luyện tự động cho các hoạt động thể thao có thể giúp người chơi nâng cao thành tích và giảm nguy cơ chấn thương. Nhiều nhà nghiên cứu đã phát triển các hệ thống vi tính hóa để đánh giá các hoạt động liên quan đến tập thể dục như xếp hạng cầu thủ bóng đá, đánh bóng ném, bóng chày, chạy nước rút, nhảy và các hoạt động thể thao khác. Patil et al. [1] đã đề xuất một dự án 'Yoga Tutor' sử dụng các đặc điểm mạnh mẽ được tăng tốc để tạo ra sự khác biệt trong các tư thế giữa một

ngư ời học và ngư ời chuyên nghiệp (SURF). Vũ và cộng sự. [2] đã trình bày một hệ thống thông minh dựa trên hình ảnh và văn bản cho yoga, như họ không xem xét tư thế của ngư ời tập. Chen và cộng sự. [3] đã sử dụng phư ơng pháp dựa trên tính năng để tạo ra một hệ thống tự đào tạo nhận dạng các bài tập yoga. Nó sử dụng Kinect để chụp đư ờng viền cơ thể của ngư ời đó và tạo bản đồ cơ thể. Để có đư ợc một mô tả cho vị trí của con ngư ời, một bộ xử lý sao đã đư ợc sử dụng để tạo khung xử lý nhanh chóng. Trong tài liệu tham khảo [4], một hệ thống nhận dạng yoga dựa trên phân loại Kinect và AdaBoost với điểm chính xác 94,78 đư ợc đề xuất cho sáu asana. Tuy nhiên, họ đang sử dụng một máy ảnh dựa trên cảm biến độ sâu mà bình thường không thể có đư ợc. Sử dụng các phư ơng pháp mạng thần kinh tích chập (CNN) và bộ mã hóa tự động xếp chồng (SAE), Mohanty et al. [5] đã triển khai phư ơng pháp nhận dạng hình ảnh để phát hiện các điệu nhảy và tư thế yoga truyền thống của Ấn Độ từ các bức ảnh. Tuy nhiên, họ chỉ phân tích năng lực của họ trên ảnh tĩnh chứ không bao giờ trên video. Kể từ khi giới thiệu DeepPose của Toshev et al. [6], phư ơng pháp khung xử lý truyền thống đã đư ợc thay thế bằng các công nghệ dựa trên học sâu. DeepPose đang dẫn đầu về các kỹ thuật truyền thống và hướng tới các phư ơng pháp tiếp cận dựa trên mạng sâu. Nó hồi quy trực tiếp trên các tọa độ chung bằng cách sử dụng các bộ hồi quy dựa trên mạng thần kinh sâu. Nó dự đoán các hoạt động của một ngư ời và cũng dự đoán vị trí của các bộ phận cơ thể bị ẩn. Tuy nhiên, cách tiếp cận của họ gặp khó khăn với nội địa hóa.

Trong những năm gần đây, có những công trình liên quan về phát hiện và phân loại tư thế yoga. [7] Các phư ơng pháp phát hiện điểm chính đư ợc sử dụng là OpenPose [8], PoseNet [9] và PifPaf [10]. Để phát hiện tư thế của con ngư ời, nhiều yếu tố sẽ đư ợc xem xét như môi trường xung quanh, tư ơng tác của con ngư ời và các biến thể của quần áo [11]. Các phư ơng pháp học sâu mà họ sử dụng để phân loại tư thế là perceptron đa lớp, mạng nơ-ron hồi quy, trí nhớ ngắn hạn dài (LSTM) [12, 13] và mạng nơ-ron tích chập. Hạn chế trong các tác phẩm trên là các tính năng (điểm chính) không đư ợc chia tỷ lệ và không thể tìm thấy mẫu cho các tư thế của con ngư ời ở các khoảng cách khác nhau so với máy ảnh. Các phư ơng pháp trước đây sử dụng các góc của khớp làm đặc điểm là bất biến xoay (ngay cả khi các khớp đư ợc xoay, các góc giữa chúng không thay đổi) [14, 15].

Trong công việc đư ợc đề xuất để phát hiện tư thế bất thường, nghiên cứu đã sử dụng các mạng phân loại tư thế yoga và tính toán độ lệch so với tư thế chuyên gia đã tính toán. dự án chủ yếu tập trung vào các bộ dữ liệu tiền xử lý để trích xuất các tính năng mới như góc giữa các bộ phận cơ thể và cách chúng cải thiện độ chính xác so với các tính năng tư thế truyền thống và bằng cách lấp đầy các giá trị còn thiếu. hệ thống sử dụng các mạng phân loại như perceptron đa lớp và điều chỉnh siêu tham số để đạt đư ợc độ chính xác cao. Giai đoạn đầu tiên của dự án nói về các công việc liên quan trong lĩnh vực này, tiếp theo là mô tả về tập dữ liệu mà nghiên cứu đã sử dụng và các kỹ thuật tiền xử lý. Làm thế nào các góc đư ợc trích xuất đư ợc thảo luận trong tiền xử lý. vi, dự án tập trung vào mã thời gian tư thế và đào tạo MLP (perceptron đa lớp) để phân loại yoga. Cuối cùng, các số liệu đánh giá mô hình và cách xây dựng các đề xuất cho ngư ời dùng sẽ đư ợc thảo luận. Tổng quan về phư ơng pháp đề xuất đư ợc thể hiện trong Hình 1.

e giấy còn lại đư ợc tổ chức như sau. Công việc liên quan đến ước tính tư thế yoga đư ợc thảo luận. bộ dữ liệu điện tử là



Hình 1: hình minh họa tổng quan về đề xuất phư ơng pháp.

phác thảo, và phần giải thích về phư ơng pháp đư ợc đề xuất để trích xuất các điểm và tính năng chính, phân loại và tạo phản hồi. Các chi tiết triển khai, các biện pháp đánh giá của phư ơng pháp đư ợc đề xuất và phân tích thời gian chạy sẽ đư ợc thảo luận. Cuối cùng, kết luận và phạm vi tư ơng lai đư ợc vạch ra.

2. Công việc liên quan

Nhận dạng hoạt động của con ngư ời đã đư ợc sử dụng trong nhiều ứng dụng, bao gồm chế tạo ngư ời máy và kỹ thuật máy tính. Tài liệu tham khảo [16, 17] sử dụng cây ngẫu nhiên (rừng ngẫu nhiên) để phát hiện các hoạt động của con ngư ời với sự trợ giúp của cảm biến. Tài liệu tham khảo [18] sử dụng các mô hình Markov ẩn và các bộ phận cơ thể đư ợc công nhận để nhận dạng hoạt động của con ngư ời. là phư ơng pháp đư ợc sử dụng để ghi nhận 6 hoạt động tại nhà, đạt độ chính xác 97,16%. là phư ơng pháp đư ợc sử dụng tại các ngôi nhà thông minh để giám sát các dịch vụ. [19] Nó sử dụng âm thanh nền môi trường để nhận dạng hoạt động của con ngư ời trong đó sử dụng cảm biến đeo đư ợc để phát hiện âm thanh, đạt độ chính xác 96,9%.

Công việc quan trọng đã đư ợc thực hiện trong việc phát triển các hệ thống tự động, phân tích các hoạt động yoga và thể thao như bóng rổ [20] và đạp xe [21]. [1] Một hệ thống tự động để ngư ời dùng ngâi thơ thực hiện yoga và so sánh với các video yoga của chuyên gia sử dụng thuật toán Tăng tốc các tính năng mạnh mẽ (SURF) chỉ sử dụng thông tin đư ờng viền, có thể không đủ.

[4] Một dự án tự động phát hiện tư thế yoga bằng cảm biến động học và bộ phân loại AdaBoost đạt độ chính xác 94,8%. Một hệ thống khác được trình bày trong tài liệu tham khảo [3] cho 3 tư thế yoga đạt độ chính xác 82,84%. [5] Hệ thống điện tử đã sử dụng các kỹ thuật học sâu để phân loại các tư thế yoga. Trong học máy truyền thống, [22] các mô hình yêu cầu các tính năng và kỹ thuật được trích xuất, nhưng học sâu hiểu dữ liệu và trích xuất các tính năng. [2] Một hệ thống tự hướng dẫn được xây dựng cho tư thế yoga bằng cách sử dụng tính toán bộ xử lý sao. Để trích xuất đường viền cơ thể từ bản đồ cơ thể người dùng, Kinect được sử dụng và đạt độ chính xác 99,33%. [23] Nó đã sử dụng phương pháp học dựa trên hàm băm để trích xuất tư thế của con người từ cảm biến áp suất. Các cảm biến này có thể không khả thi khi luôn mang theo, vì vậy trong hệ thống được đề xuất, không có cảm biến nào được sử dụng.

Ước tính tư thế được sử dụng là trong OpenPose và đã sử dụng mô hình kết hợp CNN với LSTM để phân loại các tư thế yoga và mô hình này kết hợp trích xuất tính năng. Ngoài ra, nó so sánh các mô hình CNN cơ bản với một mô hình lai và các mô hình học máy được so sánh với các mô hình học sâu. Các số liệu đánh giá được sử dụng là điểm phân loại và ma trận nhầm lẫn. SM đạt độ chính xác thử nghiệm là 0,9319, CNN đạt 0,9858 và mô hình lai CNN với LSTM đạt 0,9938. Có nhiều phương pháp phát hiện điểm chính như OpenPose, PoseNet và PifPaf. OpenPose [8] được phát minh trong kiến trúc dựa trên CMU và CNN được sử dụng để lấy các điểm chính. OpenPose sử dụng VGG-19 để trích xuất các tính năng từ hình ảnh. 18 bản đồ tin cậy đã được phát hiện bởi nhánh đầu tiên (các lớp ban đầu). nhánh thứ hai được sử dụng để dự đoán sự liên kết giữa các bộ phận cơ thể.

PoseNet [9] tương tự như OpenPose, có thể trích xuất tư thế của con người. Tất cả những điểm chính này được lập chỉ mục với mức độ tin cậy, 1 là lớn nhất và 0 là thấp nhất. PoseNet không phụ thuộc vào kích thước của hình ảnh; mặc dù hình ảnh được thu nhỏ, tư thế được trích xuất. [24] Bộ mã hóa điện tử tạo ra một vectơ mã hóa, bộ bản địa hóa chuyển đổi mã hóa thành một vectơ đặc trưng bản địa hóa và bộ hồi quy được sử dụng để hồi quy tư thế cuối cùng. PifPaf [10] dựa trên cách tiếp cận từ dưới lên để trích xuất tư thế con người. Trùng lặp cơ sở bộ phận được sử dụng để bản địa hóa các bộ phận cơ thể và Trùng lặp liên kết bộ phận được sử dụng để liên kết các bộ phận cơ thể; hai điều này kết hợp cho toàn bộ tư thế cơ thể. Kiến trúc điện tử được sử dụng là ResNet. Các tính năng được sử dụng bởi các mô hình này là 18 điểm chính; kích thước đầu vào là 36 tọa độ x và y của từng điểm chính. Các mô hình sẽ đào tạo với độ chính xác cao hơn nếu các tính năng được trích xuất từ những điểm chính này. Trong dự án, 12 tính năng cho biết các góc được trích xuất từ 12 khớp khác nhau, được sử dụng làm đầu vào cho các mô hình. Các phương pháp trước đây [25-27], sử dụng các góc nói chung đặc điểm, cho thấy các hoạt động của con người và các chuỗi chuyển động góc có liên quan với nhau. Các tính năng góc này, vì chúng có thể mở rộng, có nhiều thông tin hơn các điểm chính. Trong tài liệu tham khảo [25], người ta chỉ ra rằng các góc giữa khuỷu tay, vai, đầu gối và đáy quần đóng góp thêm thông tin để phát hiện hoạt động con người 3D. Trong tài liệu tham khảo [28], các cặp góc cho xử lý hông được thêm vào và trong tài liệu tham khảo [27], đối với hành động đứng và đi, các đặc điểm như đầu gối phải, đầu gối trái và khuỷu tay cung cấp thêm thông tin.

Trong tài liệu tham khảo [28, 29], các tính năng được sử dụng là các góc giữa các khớp sao cho các tính năng được chia tỷ lệ. Trong tài liệu tham khảo [29], các tính năng

chủ yếu được sử dụng là góc hông và đầu gối. Các góc ở khớp hông là các góc tạo bởi vai và đầu gối, và các góc ở đầu gối là các góc tạo bởi hông và mắt cá chân. các tính năng cung cấp thông tin <sup>cái đó</sup> chi tiết hơn so với các điểm chính vì ở bất kỳ khoảng cách nào so với máy ảnh, các góc được trích xuất sẽ giống nhau, nhưng không phải là các điểm chính vì chúng không được chia tỷ lệ. Trong tài liệu tham khảo [28], các góc được tính toán liên quan đến một điểm chính tham chiếu trong không gian 3D. Nhưng trong cả hai trường hợp này, các tính năng này là bất biến xoay. Trong hệ thống được đề xuất, các góc đối với trục x, nghĩa là mặt đất, được xem xét. Có 12 khớp trong đó mỗi khớp kết nối 2 điểm chính, vì vậy 12 tính năng (góc) hiện diện hoàn toàn. Giả sử a và b là hai điểm chốt thì góc tạo bởi khớp ab với trục x đóng vai trò là một đặc trưng.

3. Phương pháp luận

Trong bài báo này, một phương pháp ước tính tư thế yoga dựa trên học sâu được trình bày trong thuật toán 1 được đề xuất để phát hiện các tư thế yoga chính xác và cung cấp phản hồi để cải thiện tư thế NVIDIA phương pháp yoga e đề xuất đã được thực hiện trên DGX V-100 và bao gồm ba bước chính:

- (1) Trích xuất tính năng: video hoặc hình ảnh được cung cấp làm đầu vào cho mô hình và các khung hình được trích xuất đều đặn từ video và gửi đến ước tính tư thế nhiều người của Keras để trích xuất các điểm chính. Từ những điểm chính này, 12 vectơ chung được tính toán. Đối với tất cả 12 khớp này, lần lượt tìm được các góc giữa trục x và khớp.
- (2) Phân loại: các góc này được gửi đến mô hình phân loại để phân loại tư thế trong số 6 tư thế yoga. Các góc này được so sánh với một dãy 12 góc của tư thế đã được phân loại. là mảng chứa các góc trung bình của 12 khớp từ tập dữ liệu.
- (3) Tạo phản hồi: sự khác biệt được tính toán tương ứng cho mọi góc độ và các đề xuất được tiết lộ cho mọi góc độ. Dựa trên dấu hiệu của sự khác biệt, việc xoay các khớp theo chiều kim đồng hồ hay ngược chiều kim đồng hồ được đưa ra dự đoán đầu ra phản hồi.

Phương pháp đề xuất được trình bày dưới dạng sơ đồ trong Hình 2 và các giải thích thêm về từng bước được cung cấp trong các phần sau.

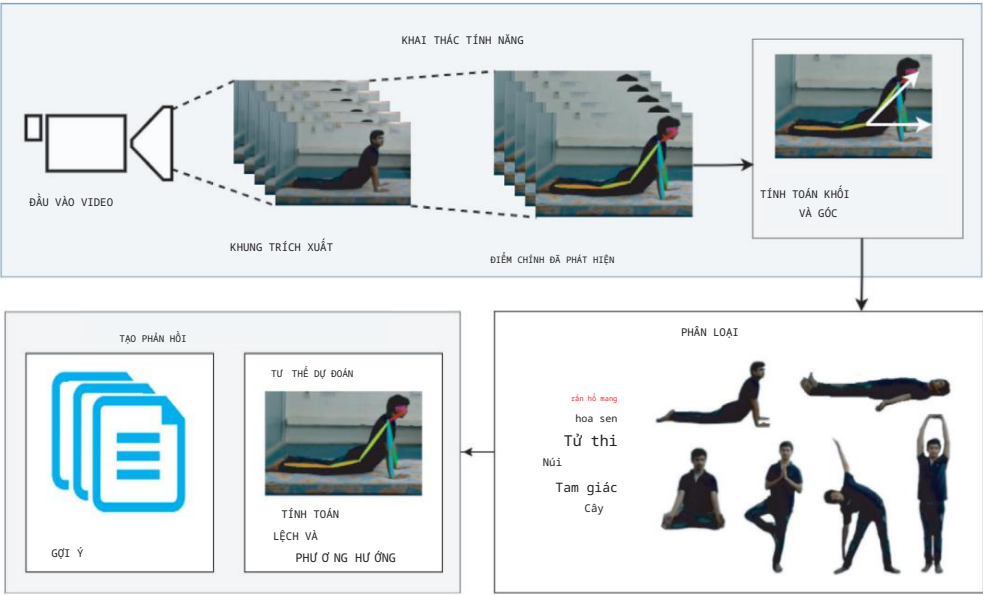
3.1. Bộ dữ liệu. Phương pháp đề xuất được kiểm tra trên bộ dữ liệu mã nguồn mở, trực tuyến, có sẵn công khai [30]. bộ dữ liệu bao gồm 6 tư thế yoga, cụ thể là Rắn hổ mang (Bhuj), Cây (Vriksh), Núi (Tada), Hoa sen (Padam), Tam giác (Trik) và Xác chết (Shav). Tổng số video của 6 tư thế là 70 và tổng số trường hợp kết hợp 6 tư thế là 350. Các video này được quay trong phòng bằng camera từ khoảng cách 4 mét; tốc độ khung hình trên giấy là 30.

Để có những người mẫu được đào tạo mạnh mẽ, các cá nhân đã thực hiện những tư thế này với một vài biến thể. Bảng 1 tóm tắt số liệu thống kê của bộ dữ liệu về số lượng video, thời lượng của từng loại hoạt động tính bằng giây và số lượng người cho từng tư thế yoga riêng biệt và một số khung mẫu của mọi tư thế được mô tả trong Hình 3.

## THUẬT TOÁN 1: Phân loại tư thế Yoga

cái đó

[illegible]



Hình 2: Một sơ đồ của phương pháp được đề xuất để ước tính tư thế yoga đúng và các thể hệ phản hồi cho tư thế không đúng.

Bảng 1: Tổng hợp số liệu thống kê của bộ dữ liệu cho từng bài yoga tư thế.

tư thế yoga	Thời gian	ngư ời	Video
tư thế rắn hổ mang	615	15	14
tư thế hoa sen	495	15	10
tư thế xác chết	450	15	10
tư thế núi	585	15	12
tư thế tam giác	540	15	13
tư thế cái cây	500	15	11
Tổng cộng			70

mức độ tự tin. Ví dụ, trong tư thế Bhuj, bên phải cổ tay, hai điểm cơ thể được phát hiện với độ tin cậy khác nhau các cấp (Hình 4).

3.3. Khai thác tính năng. Để trích xuất các điểm chính cho thời gian tạo tư thế, ước tính tư thế nhiều người trong thời gian thực của Keras là sử dụng [7, 8]. là tư thế ước tính được chạy trên mọi video, các khung hình được trích xuất cứ sau 2 giây và tư thế được tính toán cho 5 khung hình liên tiếp của mỗi video, kết quả là trong 350 trường hợp cho 70 video. Mỗi tư thế xuất ra một mảng 18 điểm chính trong đó mọi điểm bao gồm tọa độ x và y. Hình 5 cho thấy các điểm chính được trích xuất từ khung bằng mã ước tính tư thế.

Công trình nghiên cứu điện tử đã sử dụng 320 trường hợp để đào tạo. Trong khi phát hiện tư thế cho một người, nhiều điểm chính là được phát hiện với các mức tin cậy khác nhau. máy ảnh tư thế ước tính hoạt động theo cách mà nó bao gồm khóa đầu tiên điểm được phát hiện mà không tính đến độ tin cậy khoảng cách. Trong bài báo này, một vài sửa đổi đã được thực hiện đối với Keras đặt ra ước tính để xem xét các điểm chính của mức cao nhất mức độ tự tin. Với các tọa độ x và y này, nghiên cứu các tính năng được trích xuất như góc giữa các khớp cơ thể và với mặt đất để người mẫu sẽ được đào tạo để đạt được thành tích tốt sự chính xác. Ưu tiên cao nhất được dành cho các trường hợp này để

sẽ không có dữ liệu bất thường nào được đưa ra làm đầu vào. Hình 6 mô tả ước tính tư thế trên tất cả 6 tư thế yoga.

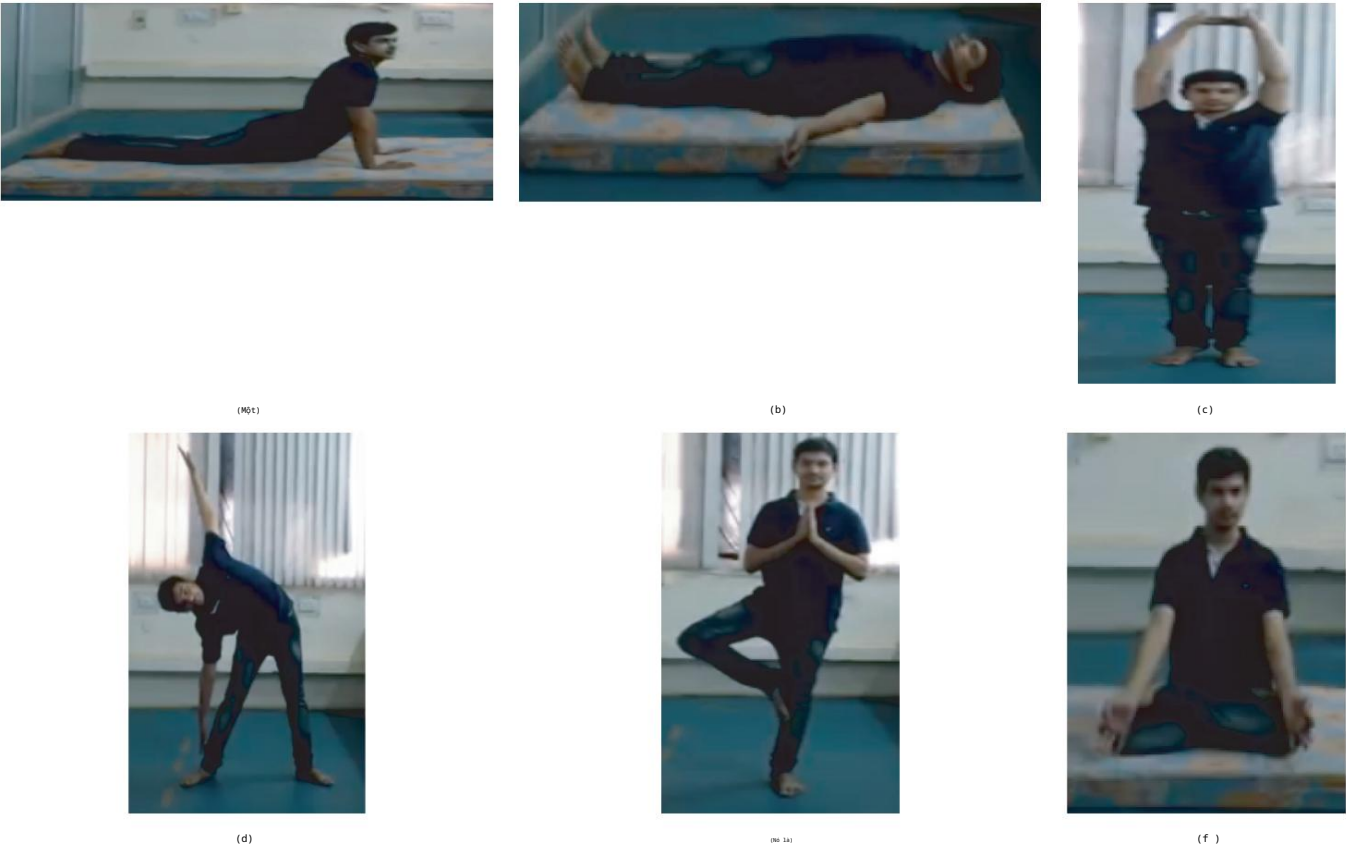
Mỗi điểm được trích xuất được coi là một kết nối vector nguồn gốc. Ở các điểm trên cơ thể, mũi, tai và mắt không có đặc điểm coi như chúng không phải là tính năng quan trọng, và các tính năng có điểm tin cậy nhỏ hơn 0,3 cũng không

xem xét để xem xét các khớp được chính xác dễ thấy. Vì vậy, số lượng vector hiện có là 13. Tổng cộng, bộ nét có 12 khớp không có mũi, tai, mắt. e 12 khớp là cổ với vai phải, vai phải với

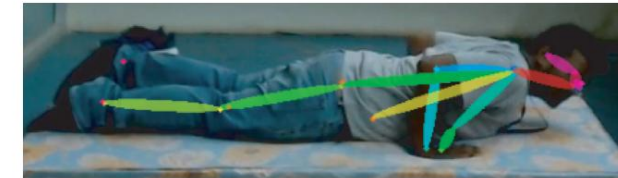
khủy tay phải, khuỷu tay phải vào cổ tay phải, cổ sang trái vai, vai trái sang khuỷu tay trái, khuỷu tay trái sang trái cổ tay, cổ đến hông phải, hông phải đến đầu gối phải, phải đầu gối sang mắt cá chân phải, cổ sang hông trái, hông trái sang trái đầu gối, và đầu gối trái đến mắt cá chân trái. Từ 13 vector này, 12

các khớp có thể thu được bằng cách trừ các vector. Giả sử, cơ thể điểm cổ và vai phải là  $(x_1, y_1)$  và  $(x_2, y_2)$ , tư ơng ứng. ví, vector của chúng là  $x_1i + y_1j$  đối với cổ và  $x_2i + y_2j$  cho vai phải. Để có được một vector cho khớp cổ và vai phải, trừ vector cổ khỏi vector vai là  $(x_2 - x_1)i + (y_2 - y_1)j$  là thể hiện trong Hình 7. Như ng, 1 nên được nhân với  $(y_2 - y_1)$  vì nguồn gốc trong hình ảnh nằm ở góc trên cùng bên trái, khác với góc dưới cùng bên trái. Vì vậy, véc tơ đối với khớp là  $(x_2 - x_1)i + (-1)(y_2 - y_1)j$ . Bằng cách này, 12 thu được các vector cho 12 khớp và các góc của chúng làm với trục x cần được tính toán. Giả sử, các góc tạo bởi một vector có trục x là theta, khi đó  $\cos(\theta)$  cho  $(x_2 - x_1)i + (-1)(y_2 - y_1)j$  là  $(x_2 - x_1) / \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$ . Với phương pháp này, thu được 12 góc cho 12 vector khác nhau cho 12 khớp. Nên bộ tính năng có 12 cột.

Các góc được trích xuất là các biến thể xoay và chia tỷ lệ. Các tư thế khác nhau khác nhau về khoảng cách từ máy ảnh cần phải chia tỷ lệ các điểm chính để huấn luyện mô hình đạt độ chính xác cao. Tuy nhiên, khi các góc được sử dụng làm đặc điểm, các khoảng cách khác nhau sẽ



Hình 3: Khung mẫu của mọi tư thế yoga: (a) Rắn hổ mang, (b) Xác chết, (c) Ngọn núi, (d) Tam giác, (e) Cái cây và (f) Tư thế hoa sen.



Hình 4: Ước tính tư thế người cho tư thế Bhuj với các mức độ tin cậy khác nhau.

không có nhu cầu mở rộng bất kỳ tính năng nào. Ví dụ: nếu khớp mắt cá chân trái và đầu gối trái tạo với mặt đất một góc 90, thì góc tạo bởi khớp này sẽ giống nhau ở mọi khoảng cách. Nếu các điểm bị xoay, giả sử khớp mắt cá chân trái và đầu gối trái xoay theo một góc nhỏ bất kỳ, thì góc tạo bởi trục x sẽ thay đổi. Trong tài liệu tham khảo [29], các góc được tạo ở hông sẽ không thay đổi nếu cả 3 điểm chính (vai, hông và đầu gối) xoay cùng một góc. Trong tham chiếu [28], nếu hai điểm chính được đổi chỗ, thì góc được tạo liên quan đến điểm tham chiếu sẽ không thay đổi. Do đó, các góc được sử dụng trong tài liệu tham khảo [28, 29] là bất biến xoay và trong hệ thống, các góc là các biến xoay.

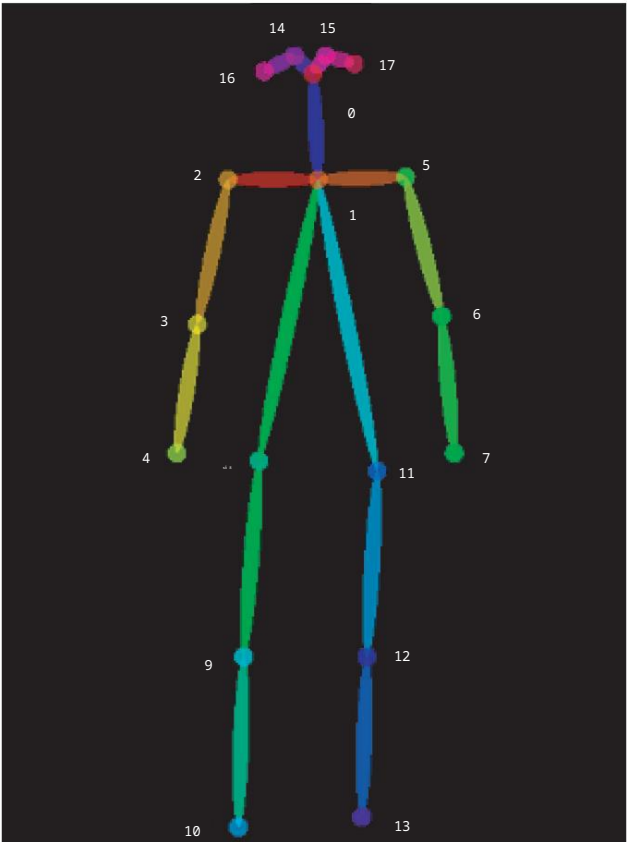
3.4. Tạo phản hồi. Trong tập dữ liệu, các giá trị hoặc góc trung bình được tính cho mọi tư thế bằng cách xem xét tất cả các tư thế được thực hiện bởi mọi người. Khi hình ảnh được cung cấp làm đầu vào cho

mô hình, mô hình được đào tạo sẽ phân loại tư thế mà nó căn chỉnh. Các góc trích xuất từ hình ảnh được so sánh với các giá trị trung bình được tính toán. Sự khác biệt giữa các góc này được tính toán, nghĩa là 12 giá trị được tính. Để đưa ra gợi ý, cần có hai tham số-tư thế bị lệch bao nhiêu so với ban đầu và theo hướng nào. Độ lớn của 12 giá trị khác biệt này cho biết mức độ một người phải sửa tư thế của mình và để biết hướng, các nhà nghiên cứu đã sử dụng dấu hiệu cho dù đó là dương hay âm, cho chúng ta biết xoay các khớp theo chiều kim đồng hồ hoặc ngược chiều kim đồng hồ. Với phương pháp này, các đề xuất được đưa ra cho người dùng cho mọi khớp.

4. Kết quả

Mạng nơ-ron (MLP) được xây dựng bằng 3 loại lớp, đó là lớp đầu vào, lớp ẩn và lớp đầu ra. Có thể là bất kỳ số lớp ẩn nào dựa trên độ phức tạp của dữ liệu huấn luyện. Nếu ít lớp ẩn, mô hình có thể không phù hợp với dữ liệu đào tạo và nếu chúng nhiều hơn, mô hình có thể quá khớp. MLP là một mạng nơ-ron được kết nối đầy đủ, nghĩa là mọi nút được kết nối với mọi nút khác trong các lớp liên tiếp trong mạng nơ-ron. Nói chung, các mạng này được sử dụng để đào tạo có giám sát trong đó đối với mọi dữ liệu đầu vào, có một nhãn hoặc lớp đầu ra tương ứng.

Perceptron đa lớp (MLP) được sử dụng để phân loại tư thế con người [31]. Trong bài báo này, các góc giữa các điểm chính đã được tính toán và chuyển thành đầu vào cho MLP. Bên trong



Hình 5: Các điểm chính được trích xuất từ khung bằng phương pháp ước tính tư thế [7].

dự án, độ dài dữ liệu đầu vào là 12 và có 6 lớp để phân loại các nhân này, vì vậy độ dài của lớp đầu ra là 6. Hình 8 mô tả kích thước lớp đầu vào là 12, kích thước lớp ẩn đầu tiên là 10, kích thước lớp ẩn thứ 2 là 8 và kích thước lớp đầu ra là 6. Tổng cộng, có 350 trữ ứng hợp, để huấn luyện có 320 trữ ứng hợp được sử dụng, trong khi đối với mỗi tư thế, 5 trữ ứng hợp được sử dụng để xác nhận. kích thước lô được sử dụng để đào tạo là 20 và số kỷ nguyên là 10000. Hình 9 cho thấy đồ thị của độ chính xác và mất mát cho tập dữ liệu đào tạo và xác nhận.

Cả tập dữ liệu đào tạo và xác nhận đều có nhiều ưu điểm và giảm độ chính xác cho đến kỷ nguyên 6900 và đạt được độ chính xác là 0,9958 ở kỷ nguyên 6900. Từ 6900 đến 10000 kỷ nguyên, mất đào tạo và xác nhận giảm dần, dẫn đến mô hình đào tạo được phân loại cao sự tự tin. Từ kỷ nguyên 0 đến 10000, việc mất xác thực và tập dữ liệu huấn luyện giảm dần. Từ việc đào tạo và độ chính xác xác nhận thu được, có thể suy ra rằng mô hình không quá phù hợp. chức năng mất e được sử dụng là phân loại entropy chéo vì nghiên cứu đang phân loại các tính năng đầu vào vào một trong 6 nhân. Trình tối ưu hóa AdaDelta được sử dụng dựa trên về tỷ lệ học tập thích ứng để giải quyết hai nhược điểm: (1) phân rã tỷ lệ học tập và (2) lựa chọn toàn cầu tỷ lệ học. Chức năng kích hoạt e được sử dụng cho lớp cuối cùng là

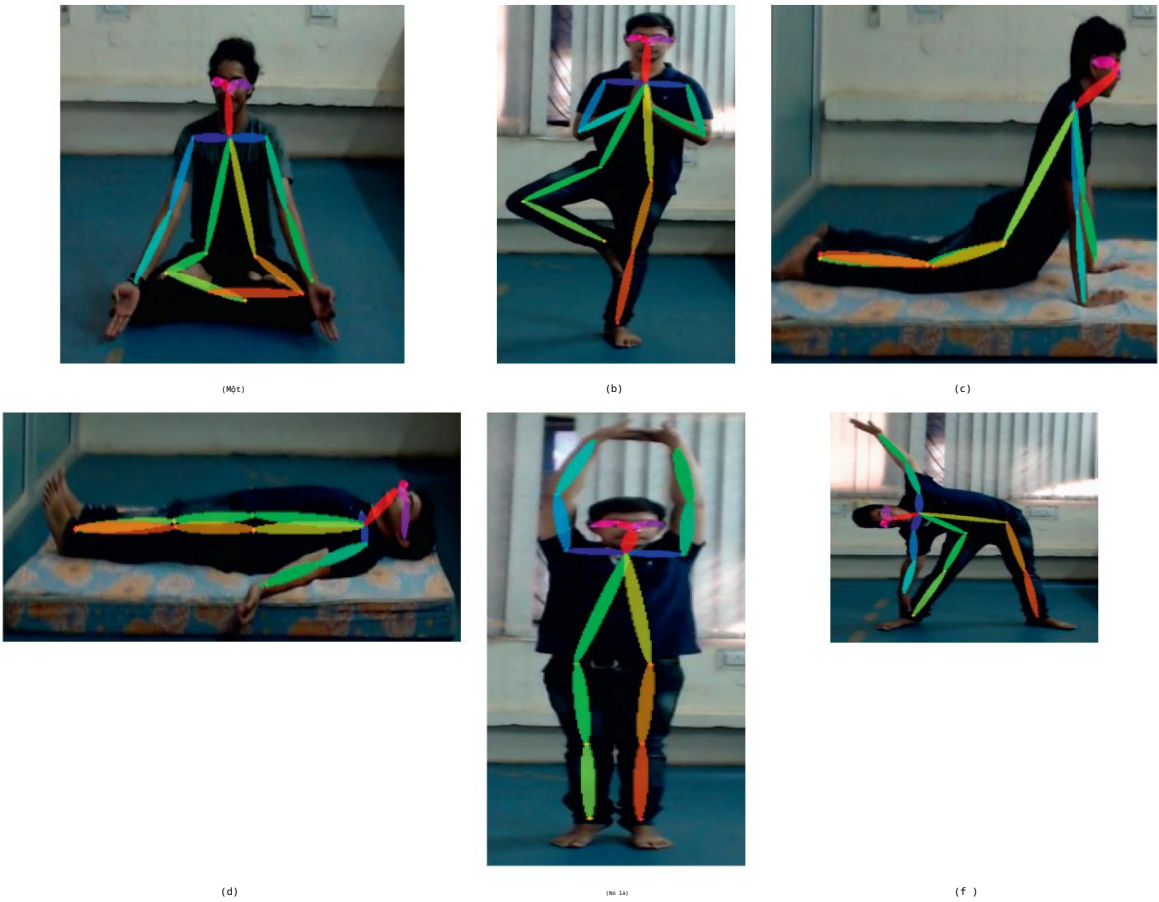
softmax vì nó đưa ra mức độ tin cậy cho tất cả các nhân. một với độ tin cậy cao nhất là nhân dự đoán.

Bảng 2 thể hiện kết quả độ chính xác của các mô hình thực nghiệm, SVM thu được kết quả độ chính xác là 0,9319, CNN thu được kết quả chính xác là 0,9858 và CNN + LSTM đạt kết quả chính xác 0,9938. sức mạnh MLP trong hệ thống nhỏ hơn đáng kể so với CNN và CNN + LSTM, nhưng nó thu được độ chính xác 0,9958 với các tính năng được sửa đổi.

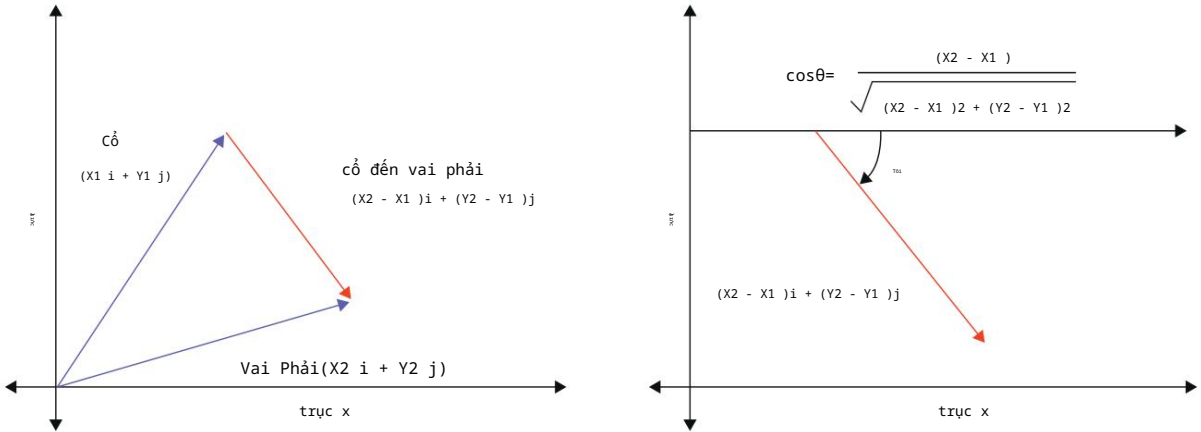
Để kiểm tra tính hiệu quả của phương pháp luận được đề xuất, một ma trận nhầm lẫn được sử dụng mô tả hiệu suất mô hình phân loại cho tất cả các trữ ứng hợp về mặt độ chính xác theo phương trình (1). Độ chính xác phân loại cũng được gọi là điểm phân loại, là tỷ lệ đúng phân loại và tổng số trữ ứng hợp.

ma trận nhầm lẫn được sử dụng trong nghiên cứu có 6 nhân, vì vậy đánh giá kết quả có ma trận nhầm lẫn 6 x 6. nó \_ hàng đại diện cho lớp thực tế, trong khi đại diện cột thứ j gửi lại lớp dự đoán của dữ liệu được đề xuất. Hình 10 mô tả các ma trận nhầm lẫn của đào tạo, xác nhận và bộ dữ liệu thử nghiệm. Trong ma trận nhầm lẫn của đào tạo, xác thực và tập dữ liệu huấn luyện, tổng số vị trí trong lần lượt là 320, 30 và 30. Nó có thể quan sát rằng tất cả các mẫu được dự đoán chính xác, dẫn đến kết quả là





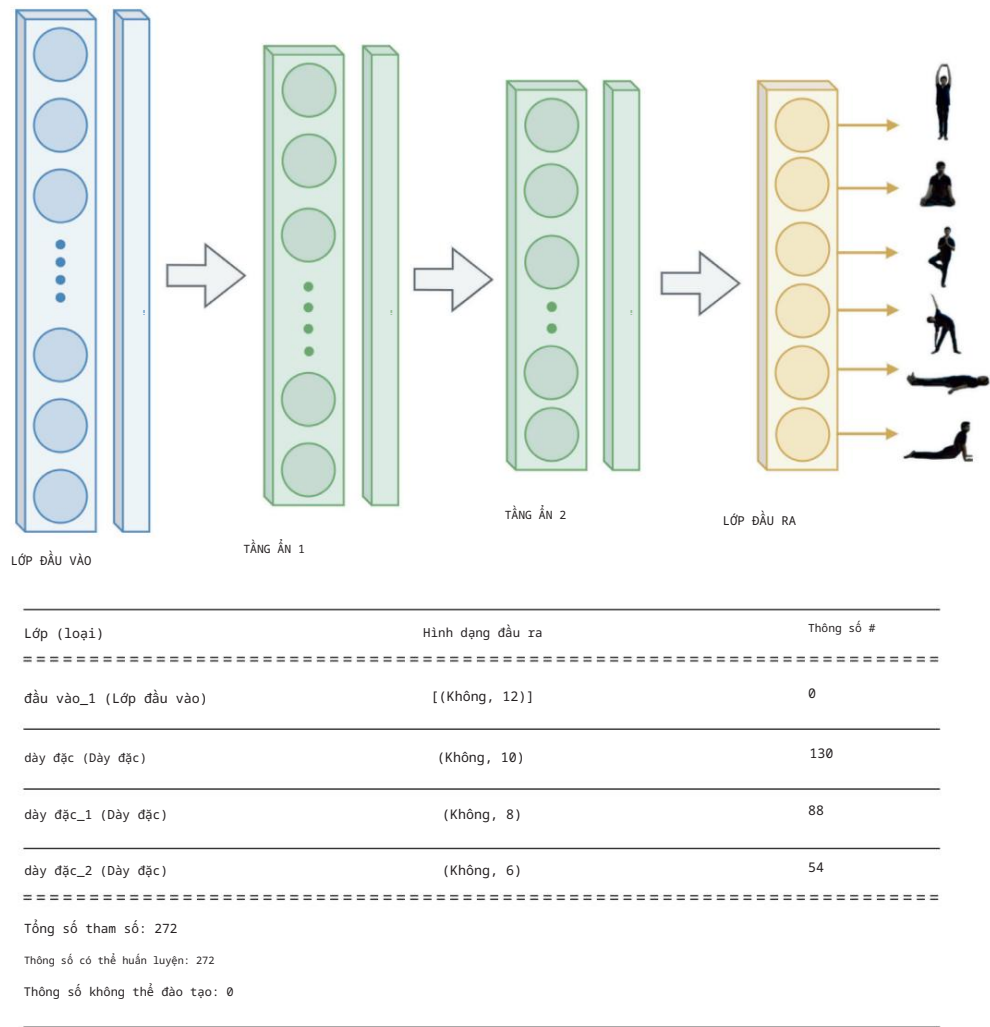
Hình 6: Minh họa khai thác các điểm chính trên tất cả 6 tư thế yoga: (a) Tư thế hoa sen, (b) Tư thế cái cây, (c) Tư thế rắn hổ mang, (d) Tư thế xác chết, (f) Tư thế tam giác và (e) Tư thế ngọn núi.



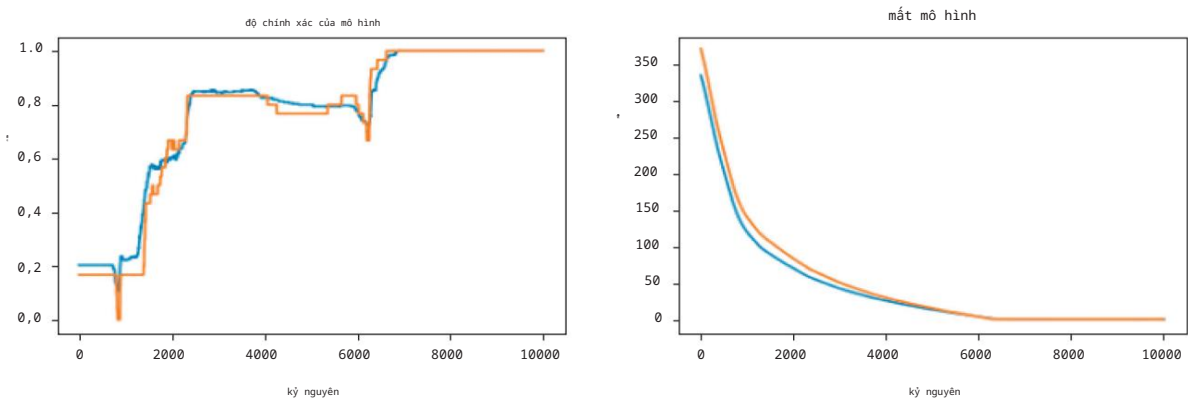
Hình 7: e biểu diễn véc tơ tọa độ và tính góc của véc tơ cổ và vai phải.

độ chính xác 0,9958 cho mọi trường hợp. Hình 11 thể hiện đồ thị cho các mô hình cạnh tranh khác nhau.





Hình 8: Kiến trúc mô hình mạng nơ-ron.



Hình 9: Đồ thị về độ chính xác và tổn thất cho tập dữ liệu huấn luyện và xác thực.

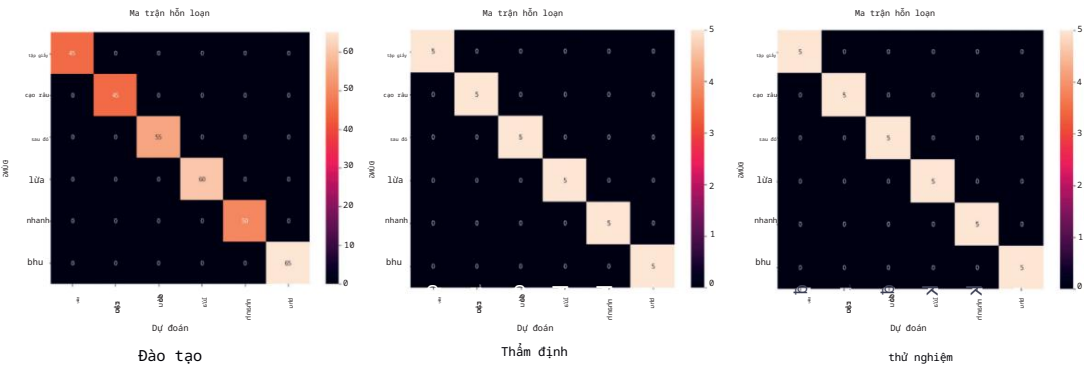
5. Phân tích thời gian chạy

phương pháp tiếp cận được trình bày trong nghiên cứu này được dựa trên học sâu để phát hiện tư thế yoga không chính xác và cũng khuyến khích người dùng cải thiện tư thế. Trong nghiên cứu này,

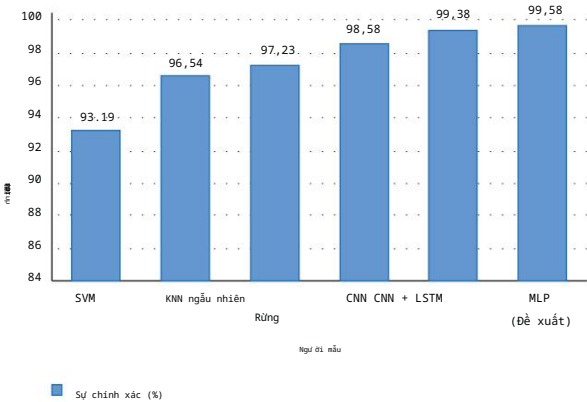
trích xuất các điểm chính bằng cách sử dụng kỹ thuật ước tính tư thế, tính toán các vectơ cho mỗi khớp và góc giữa các vectơ cho các khớp liên kế được phân loại là đặc trưng. Sau đó, các tính năng này đã được đưa vào kỹ thuật phân loại, và sau đó, thông tin phản hồi cho

Bảng 2: Bảng thể hiện kết quả độ chính xác của các mô hình thí nghiệm.

Người mẫu	Sự chính xác	
	Đào tạo	thử nghiệm
SVM	0,9532	0,9319
CNN	0,9934	0,9858
CNN + LSTM	0,9987	0,9938
MLP	0,9962	0,9958



Hình 10: Ma trận nhầm lẫn của tập dữ liệu đào tạo, xác nhận và kiểm tra. (a) Đào tạo, (b) xác nhận và (c) thử nghiệm.



Hình 11: đồ thị diện tử minh họa đồ thị của các mô hình cạnh tranh khác nhau.

Bảng 3: Đánh giá thời gian chạy công việc được đề xuất tính bằng mili giây với các kỹ thuật khác nhau đang chạy. Mean (FE+C) thể hiện giá trị trung bình thời gian chạy trung bình trên mỗi khung để trích xuất và tính toán các tính năng với phân loại tư thế yoga. Mean (FG) là thời gian chạy trung bình trung bình mỗi khung để tạo phản hồi.

phương pháp	CPU		GTX-1080	
	Trung bình (FE+C)	Trung bình (FG)	Trung bình (FE+C)	Trung bình (FG)
SVM	6574,8 ± 134,7	16.4	214,5 ± 10,6	10.32
CNN	6528,6 ± 112,6	15,2	208,4 ± 15,2	8,36
CNN + LSTM	6512,6 ± 118,4	16,6	206,8 ± 16,8	7.18
MLP	6504,4 ± 124,3	12.3	206,3 ± 12,6	6,47

tính đúng đắn của tư thế yoga được tạo ra. do đó, các thời gian chạy được chia thành ba phần: (1) trích xuất và tính toán thời gian của các tính năng cho mọi khung hình, (2) phân loại và (3) thời gian tạo phản hồi của phân loại tư thế yoga trên mỗi khung hình. e thời gian chạy cho quá trình trích xuất và

tính toán của các tính năng không đổi cho mỗi phương pháp. phân tích thời gian chạy điện tử được thực hiện trên CPU Xeon(R) E3-1240 v5 và NVIDIA GeForce GTX-1080. Bảng 3 trình bày thời gian chạy trung bình trung bình trên mỗi khung hình cùng với độ lệch chuẩn của thí nghiệm

phương pháp luận. thời gian được trình bày bằng mili giây. Nó trong các công ty là thời gian dành cho mỗi khung hình khi trích xuất và phân loại tính năng với sự phát triển phản hồi.

6. Kết luận

Các phương pháp tiếp cận được trình bày trong nghiên cứu này dựa trên học sâu để phát hiện tư thế yoga không đúng và cũng khuyến khích người dùng cải thiện tư thế bằng cách xác định vị trí sai của tư thế yoga. Trong hệ thống được đề xuất, người dùng có thể chọn tư thế mong muốn để luyện tập và có thể tải lên các video đã ghi về tư thế luyện tập yoga của họ. Nghiên cứu điện tử đã trích xuất các góc hoạt động giám sát và sử dụng chúng như một tính năng khi chúng được thu nhỏ. Trong một số trường hợp, nếu các điểm chính bị xoay thì các góc không thay đổi, điều này không mang lại kết quả tốt. Trong hệ thống này, các góc với mặt đất được xem xét chứ không phải giữa các khớp, vì vậy nếu có bất kỳ sự xoay nhẹ nào của các điểm chính, thì các góc đó sẽ bị thay đổi. Với các tính năng này, perceptron đa lớp được đào tạo để đạt được độ chính xác 0,9958 cho các bộ dữ liệu thử nghiệm. Trong tìm kiếm lại hiện có, SVM đạt được độ chính xác thử nghiệm là 0,9319, CNN đạt được 0,9858 và CNN + LSTM đạt được 0,9938. Trong hệ thống, công suất MLP thấp hơn nhiều so với CNN và CNN + LSTM nhưng đạt độ chính xác 0,9958 với các tính năng được sửa đổi. Khi so sánh với các kỹ thuật hiện có, kết quả thử nghiệm cho thấy kết quả đầy hứa hẹn. phương pháp đề xuất duy trì độ phức tạp tính toán thấp, có thể được áp dụng cho cuộc sống bận rộn của ai đó để tự học yoga và có thể phát hiện tư thế yoga không chính xác để tránh các vấn đề mãn tính.

Phạm vi tương lai. Hệ thống đề xuất bị giới hạn trong 6.6.1. tư thế yoga, trong đó có tổng cộng hơn 80 tư thế yoga. Tập dữ liệu được đề xuất có thể được mở rộng bằng cách thêm các điểm chính của tư thế yoga cần thiết. Công nghệ điện tử cũng có thể được sử dụng để đưa ra dự đoán theo thời gian thực và tự đào tạo trên thiết bị di động. Có một số trường hợp ứng dụng thực tế trong đó việc đánh giá một tư thế riêng lẻ sẽ không đủ; ví dụ: ước tính tư thế trong môi trường đông đúc sẽ cần phát hiện và nhận ra tư thế của từng người tham gia. Để bao gồm nhiều tư thế và để người mẫu làm việc trên nhiều tư thế (phân loại nhiều tư thế) là đủ thách thức. Keras đặt ra ước tính ảnh hưởng đến hiệu suất của mô hình; nên thực hiện các bước lấy điểm chính khi các phần cơ thể bị trùng hoặc thiếu để đạt kết quả tốt hơn. là phương pháp trích xuất các góc vì các tính năng có thể được sử dụng cho các ứng dụng khác như phát hiện hoạt động và theo dõi hoạt động thể thao.

Dữ liệu sẵn có

dữ liệu hỗ trợ các phát hiện có sẵn theo yêu cầu tác giả tương ứng.

Xung đột lợi ích

Các tác giả của bản thảo này tuyên bố rằng không có xung đột lợi ích.

Người giới thiệu

[1] S. Patil, A. Pawar, A. Peshave, AN Ansari và A. Navada, "Thực quan hóa và phân tích người hướng dẫn yoga bằng thuật toán SURF," trong Kỷ yếu của Hội thảo nghiên cứu sau đại học về hệ thống và điều khiển của IEEE năm 2011, trang 43 -46, IEEE, Shah Alam, Malaysia, tháng 6 năm 2011.

[2] W. Wu, W. Yin và F. Guo, "Hệ thống chuyên gia tự học và tự hướng dẫn cho Yoga," trong Kỷ yếu Hội thảo quốc tế lần thứ 2 năm 2010 về Hệ thống và ứng dụng thông minh, trang 1-4, IEEE , Mumbai, Ấn Độ, tháng 5 năm 2010.

[3] HT Chen, YZ He, CL Chou, SY Lee, BSP Lin và JY Yu, "Hệ thống tự đào tạo có sự hỗ trợ của máy tính dành cho các bài tập thể thao sử dụng động cơ," trong Kỷ yếu của Hội nghị Quốc tế IEEE 2013 về Đa phương tiện và Triết lý Hội thảo (ICMEW), trang 1-4, IEEE, London, UK, tháng 7 năm 2013.

[4] EW Trejo và P. Yuan, "Nhận dạng tư thế Yoga thông qua hệ thống tư duy tác với thiết bị Kinect," trong Kỷ yếu Hội nghị Quốc tế lần thứ 2 năm 2018 về Khoa học Robot và Tự động hóa (ICRAS), trang 1-5, IEEE, Vũ Hán, Trung Quốc, tháng 6 năm 2018.

[5] A. Mohanty, A. Ahmed, T. Goswami, A. Das, P. Vaishnavi và RR Sahay, "Nhận dạng tư thế mạnh mẽ bằng cách sử dụng học sâu," trong Kỷ yếu của Hội nghị Quốc tế về Thị giác Máy tính và Xử lý Hình ảnh, trang .93-105, Springer, Singapore, tháng 12 năm 2017.

[6] A. Toshev và C. Szegedy, "Chụp sâu: ước tính tư thế con người thông qua mạng thần kinh sâu," trong Kỷ yếu hội nghị IEEE về thị giác máy tính và nhận dạng mẫu, trang 1653-1660, Columbus, OH, Hoa Kỳ, tháng 8 2014.

[7] Z. Cao, T. Simon, SE Wei và Y. Sheikh, "Ước tính tư thế 2d nhiều người trong thời gian thực bằng cách sử dụng các trường ánh lực từng phần," trong Kỷ yếu hội nghị IEEE về thị giác máy tính và nhận dạng mẫu, trang 7291-7299 , Honolulu, HI, Hoa Kỳ, Tháng 7 2017.

[8] Z. Cao, G. Hidalgo, T. Simon, SE Wei và Y. Sheikh, "OpenPose: ước tính tư thế 2D nhiều người trong thời gian thực bằng cách sử dụng Trường sở thích một phần," Giao dịch của IEEE về Phân tích mẫu và Trí tuệ máy, tập. 43, không. 1, trang 172-186, 2019.

[9] A. Kendall, M. Grimes, và R. Cipolla, "Posenet: mạng tích chập dành cho tái định vị máy ảnh 6-dof thời gian thực," trong Kỷ yếu của hội nghị quốc tế IEEE về thị giác máy tính, trang 2938- 2946, Santiago, Chile, tháng 12 năm 2015.

[10] S. Kreiss, L. Bertoni và A. Alahi, "Pifpaf: các trường tổng hợp để ước tính tư thế con người," trong Kỷ yếu của Hội nghị IEEE/CVF về Tầm nhìn Máy tính và Nhận dạng Mẫu, trang 11977-11986, Long Beach , CA, Hoa Kỳ, tháng 6 năm 2019.

[11] W. Gong, X. Zhang, J. Gonzalez và cộng sự, "Đánh giá tư thế con người từ hình ảnh một mắt: một cuộc khảo sát toàn diện," Cẩm biển, tập. 16, không. 12, tr. 1966, 2016.

[12] K. Pothanackicker, "Nhận dạng hành động của con người bằng cách sử dụng CNN và LSTM-RNN với mô hình chú ý," Tạp chí Công nghệ Sáng tạo Quốc tế. và Khám phá tiếng Anh, tập. 8, không. 8, 2019.

[13] RA Güler, N. Neverova và I. Kokkinos, "Densepose: ước tính tư thế con người dày đặc trong tự nhiên," trong Kỷ yếu của Hội nghị IEEE về Tầm nhìn Máy tính và Nhận dạng Mẫu, trang 7297-7306, Salt Lake Thành phố, UT, Hoa Kỳ, tháng 6 năm 2018.

[14] H.-T. Chen, Y.-Z. Anh ấy, và C.-C. Hsu, "Hệ thống đào tạo yoga có sự hỗ trợ của máy tính," Công cụ và ứng dụng đa phương tiện, tập. 77, không. 18, trang 23969-23991, 2018.

[15] S. Jain, A. Rustagi, S. Saurav, R. Saini và S. Singh, "kiến trúc học sâu lấy cảm hứng từ CNN 3 chiều để nhận dạng tư thế Yoga trong môi trường thể giới thực," Neural

Máy tính & Ứng dụng, tập. 33, không. 12, trang 6427-6441, 2021.

[16] CC Hsieh, BS Wu và CC Lee, “Hệ thống học yoga hỗ trợ thị giác máy tính từ xa,” Tạp chí Máy tính, tập. 6, không. 11, trang 2382-2388, 2011.

[17] MT Uddin và MA Uddiny, “Nhận dạng hoạt động của con người từ các cảm biến đeo được bằng cách sử dụng các cây cực kỳ ngẫu nhiên,” trong Kỷ yếu Hội nghị Quốc tế 2015 về Kỹ thuật Điện và Công nghệ Truyền thông Thông tin (ICEEICT), trang 1-6, IEEE, London, Vương quốc Anh, tháng 5 năm 2015.

[18] A. Jalal, N. Sarif, JT Kim, và T.-S. Kim, “Nhận dạng hoạt động của con người thông qua các bộ phận cơ thể được công nhận của bóng người đi sâu cho các dịch vụ giám sát cư dân tại nhà thông minh,” Môi trường trong nhà và xây dựng, tập. 22, không. 1, trang 271-279, 2013.

[19] Y. Zhan và T. Kuroda, “Nhận dạng hoạt động của con người dựa trên cảm biến có thể đeo được từ âm thanh nền môi trường,” Tạp chí Ambient Intelligence and Humanized Computing, tập. 5, không. 1, trang 77-89, 2014.

[20] P.-F. Pai, L.-H. ChangLiao, và K.-P. Lin, “Phân tích trò chơi bóng rổ bằng máy vectơ hỗ trợ với mô hình cây quyết định,” Ứng dụng & Điện toán Thần kinh, tập. 28, không. 12, trang 4159-4167, 2017.

[21] S. Haque, ASA Rabby, MA Laboni, N. Neehal và SA Hossain, “ExNET: mạng lưới thần kinh sâu để phát hiện tư thế tập thể dục,” trong Kỷ yếu Hội nghị Quốc tế về Xu hướng Gần đây trong Xử lý Hình ảnh và Nhận dạng Mẫu, trang 186-193, Springer, Singapore, tháng 12 năm 2018.

[22] J. Palanimeera và K. Ponmozhi, “Phân loại tư thế yoga bằng kỹ thuật máy học,” Materials Today Proceedings, vol. 37, trang 2930-2933, 2021.

[23] L. Casas, N. Navab và S. Demirci, “Ước tính tư thế cơ thể 3D của bệnh nhân từ hình ảnh áp lực,” Tạp chí Quốc tế về X quang và Phẫu thuật có sự hỗ trợ của máy tính, tập. 14, không. 3, trang 517-524, 2019.

[24] Y. Shavit và R. Ferens, “Giới thiệu về ước tính tư thế máy ảnh với học sâu,” 2019, <https://arxiv.org/abs/1907.05272>.

[25] ND ang, T.-S. Kim, Y.-K. Lee và S. Lee, “Ước tính tư thế cơ thể người đi 3-D thông qua đồng đăng ký mô hình người đi 3-D và thông tin âm thanh nổi tuần tự,” Intelligence ứng dụng, tập. 35, không. 2, trang 163-177, 2011.

[26] MZ Uddin, ND ang và TS Kim, “Nhận dạng hoạt động của con người thông qua các tính năng góc khớp 3-D và các mô hình Markov ẩn,” trong Kỷ yếu Hội nghị Quốc tế IEEE 2010 về Xử lý Hình ảnh, trang 713-716, IEEE, Hồng Kông, Trung Quốc, tháng 9 năm 2010.

[27] F. Ofli, R. Chaudhry, G. Kurillo, R. Vidal, và R. Bajcsy và Đại diện hình ảnh, tập. 25, không. 1, tr. 24-38,

..

Yoon, Ince, IF Ince, ME Yildirim, JS Park, JK Song, và [28] OF BW “Nhận dạng hoạt động của con người bằng phân tích các góc giữa các khớp xương bằng cảm biến độ sâu RGB,” Tạp chí ETRI, tập. 42, không. 1, tr. 78-89, 2020.

[29] A. Guler, N. Kardaris, S. Chandra và cộng sự, “Ước tính góc khớp của con người và nhận dạng cử chỉ đối với tầm nhìn rô-bốt hỗ trợ,” trong Kỷ yếu của Hội nghị Châu Âu về Tầm nhìn Máy tính, trang 415-431, Springer, Amsterdam, Hà Lan, tháng 10 năm 2016.

[30] SK Yadav, A. Singh, A. Gupta và JL Raheja, “Nhận dạng Yoga trong thời gian thực bằng cách sử dụng học sâu,” Ứng dụng & Điện toán Thần kinh, tập. 31, không. 12, trang 9349-9361, 2019.

[31] P. Szczuko, “Mạng lưới thần kinh sâu để ước tính tư thế con người từ một hình ảnh độ sâu có độ phân giải rất thấp,” Công cụ và Ứng dụng Đa phương tiện, tập. 78, không. 20, trang 29357-29377, 2019.