

Chỉnh sửa và ước tính các tư thế Yoga bằng cách sử dụng Mở Kiến trúc CV và VGG 19

Tanmay Hande¹
Pranali Dhawas²

Bhargavi Kakirwar³
Aaditya Gupta⁴

Cao đẳng Kỹ thuật GH Raisoni, Nagpur, Ấn Độ

Tóm tắt: - Đại dịch COVID-19 đã thay đổi đáng kể cách chúng ta tập thể dục, với nhiều người chuyển sang tập thể dục tại nhà như một cách để duy trì hoạt động trong thời gian yêu cầu ở nhà. Tuy nhiên, nếu không được tiếp cận với các huấn luyện viên chuyên nghiệp, những người mới bắt đầu có thể gặp khó khăn trong việc thực hiện các bài tập với hình thức phù hợp, làm tăng nguy cơ chấn thương. Do đó, cần có các hệ thống theo dõi hiệu suất tập thể dục để ngăn ngừa chấn thương ngắn hạn và dài hạn. Trong nghiên cứu này, chúng tôi trình bày một cách tiếp cận để phát hiện và điều chỉnh chính xác các tư thế yoga bằng cách sử dụng các kỹ thuật ước tính tư thế với kiến trúc OpenCV và VGG-19 với học chuyển GPU. Để đo và điều chỉnh chính xác tư thế cơ thể trong các buổi tập, giải pháp được đề xuất kết hợp các thuật toán dựa trên học sâu và phương pháp tiếp cận thị giác máy tính. Để xác nhận tính hiệu quả của mô hình VGG-19 trên tập dữ liệu được sử dụng, chúng tôi đã tiến hành một số lượng lớn các thử nghiệm, so sánh hiệu suất của một số chiến lược học máy và học sâu để ước tính các tư thế yoga. Với độ chính xác 98,11%, kết quả cho thấy tính hữu ích của kỹ thuật được đề xuất trong việc nhận biết và điều chỉnh chính xác các tư thế tập thể dục.

Những phát hiện của nghiên cứu này có ý nghĩa quan trọng trong việc cải thiện hiệu quả và sự an toàn của yoga các phiên và có thể được mở rộng sang các miền khác yêu cầu ước tính tư thế con người chính xác.

TÓM TẮT GIỚI THIỆU

Sự phổ biến của yoga đã tăng lên trong những năm gần đây, với nhiều người kết hợp nó vào thói quen hàng ngày của họ vì nhiều lợi ích về sức khỏe thể chất và tinh thần. Tuy nhiên, tư thế và hơi thở không đúng trong các buổi tập yoga có thể dẫn đến đau dữ dội và các vấn đề mãn tính, điều này làm nổi bật nhu cầu về công nghệ có thể giúp các cá nhân điều chỉnh và cải thiện tư thế yoga của họ [1]. Trong số những kỹ thuật này, các tư thế thể chất, được gọi là asana, đã trở nên đặc biệt phổ biến ở thế giới phương Tây. Khi đại dịch tiếp tục diễn ra, nhiều người đã chuyển sang tập yoga như một phương tiện để giữ cho mình khỏe mạnh về thể chất và tinh thần. Tuy nhiên, điều quan trọng là phải thực hiện các asana đúng cách, vì duỗi không đúng cách hoặc thực hiện các asana không phù hợp và thở không đúng cách có thể dẫn đến đau dữ dội và các vấn đề mãn tính. Điều này nhấn mạnh nhu cầu ngày càng tăng đối với các công cụ và công nghệ có thể giúp các cá nhân điều chỉnh và cải thiện các tư thế yoga của họ [2]. Để giải quyết nhu cầu này, một phân tích khoa học về thực hành asana đã

được phát triển để đảm bảo thực hành yoga an toàn và hiệu quả.

Một trong những công nghệ đang ngày càng trở nên phổ biến là việc sử dụng ước tính tư thế bằng trí tuệ nhân tạo (AI). Ước tính tư thế là quá trình sử dụng các thuật toán thị giác máy tính để theo dõi và phân tích các chuyển động của cơ thể một cá nhân. Bằng cách sử dụng công nghệ này, có thể ước tính vị trí của các khớp và bộ phận cơ thể khác nhau trong thời gian thực và cung cấp phản hồi về tính đúng đắn của tư thế yoga [3]. Trong nghiên cứu này, việc sử dụng ước tính tư thế với AI có một số lợi thế để điều chỉnh và ước tính các tư thế yoga. Thứ nhất, nó cung cấp một đánh giá khách quan và chính xác hơn về tư thế, so với việc dựa vào sự quan sát chủ quan của con người.

Thứ hai, nó cho phép phản hồi theo thời gian thực, điều này có thể cực kỳ hữu ích cho những người đang học các tư thế yoga hoặc đang cố gắng cải thiện hình thức của họ. Cuối cùng, nó có thể là một công cụ hữu ích cho các giáo viên yoga muốn cung cấp hướng dẫn và hỗ trợ được cá nhân hóa hơn cho học viên của họ.

Khả năng thực hiện tối ưu hóa từ đầu đến cuối của các kỹ thuật học sâu là một trong những lợi ích chính của nó [4]. Ước tính tư thế và nhận dạng hành động thường khó tích hợp để thực hiện tối ưu hóa chung hữu ích, đôi khi yêu cầu chuyển đổi 3D hoặc chuyển đổi bản đồ nhiệt [5]. Chuỗi lan truyền ngược cần thiết cho quá trình học từ đầu đến cuối bị gián đoạn khi sử dụng các kỹ thuật dựa trên phát hiện, vì chúng cần hàm argmax không thể phân biệt làm bước xử lý sau để khôi phục tọa độ chung. Để giải quyết vấn đề này, chúng tôi cung cấp một phương pháp mới để ước tính tư thế 2D và 3D kết hợp dựa trên phần mở rộng của soft-argmax khả vi [6]. Do đó, chúng ta có thể xây dựng một khung đa tác vụ hoàn toàn có thể đào tạo được bằng cách xếp lớp nhận dạng hành động lên trên ước tính tư thế. Tuy nhiên, việc không có các công cụ và huấn luyện viên nói trên có thể là một trở ngại đáng kể đối với việc tập yoga tại nhà của chúng ta. Mục tiêu của chúng tôi là tạo ra một huấn luyện viên dựa trên AI cho phép bạn tập yoga tại nhà hiệu quả hơn. Mục tiêu của nghiên cứu này là tạo ra một hệ thống AI hỗ trợ tập thể dục bằng cách ước tính tư thế để xác định chất lượng và số lần lặp lại. Nghiên cứu này nhằm mục đích làm cho việc tập thể dục trở nên dễ dàng và thú vị hơn, bao gồm một giao diện không phá hủy.

Có rất nhiều ứng dụng trên thị trường hướng dẫn người dùng hoàn thành các bài tập [7]. Tuy nhiên, với ứng dụng của chúng tôi, chúng tôi không chỉ hướng dẫn người dùng nên thực hiện bài tập nào mà còn hướng dẫn cách giữ mình đúng cách và

sử dụng thị giác máy tính để đếm số lần lặp lại. Chương trình này có thể được coi là một người trợ giúp yoga cung cấp khả năng phát hiện tư thế tức thì và tư vấn chế độ ăn uống. Bằng cách mở rộng phạm vi sử dụng, chương trình có thể được sử dụng tại các phòng tập thể dục với tư cách là huấn luyện viên thông minh, giảm nhu cầu tham gia của con người trong khi vẫn cho phép người dùng sử dụng tại nhà.

Trong công việc của mình, chúng tôi trình bày một ứng dụng thể dục được hỗ trợ bởi AI có thể nhận dạng tư thế yoga của người dùng, ghi lại số lần lặp lại bài tập theo quy định và cung cấp thông tin chuyên sâu, được cá nhân hóa về cách cải thiện tư thế cơ thể của người dùng. Để giúp những người không có điều kiện đến phòng tập thể dục nhưng vẫn sẵn sàng tập luyện tại nhà để duy trì vóc dáng và thể lực cũng như giữ cho cơ thể luôn ở trạng thái tuyệt vời, một hướng dẫn tập thể dục và trợ lý tập thể dục dựa trên AI đã được phát triển. để hỗ trợ họ thực hiện các bài tập đúng cách và giúp họ không bị bất kỳ chấn thương ngắn hạn hoặc dài hạn nào. Cùng với lượng calo yoga hàng ngày được cá nhân hóa, điều này cũng cung cấp lời khuyên về sức khỏe và kế hoạch ăn uống được cá nhân hóa. Các phần tiếp theo của bài tiểu luận này được cấu trúc như sau: Các chiến lược và công cụ về phương pháp luận được nêu trong Phần II. Trong Phần III, chúng tôi mô tả những phát hiện từ cuộc điều tra được đề xuất. Phần IV thảo luận về kết quả và những việc cần làm tiếp theo.

II. KHẢO SÁT VĂN HỌC

Việc phát triển các hệ thống ước tính tư thế có thể nhận biết và theo dõi các vị trí cơ thể con người từ ảnh hoặc video đã trở nên khả thi nhờ những cải tiến về thị giác máy tính và máy học. Những phương pháp này đã cho thấy thành công trong nhiều bối cảnh, bao gồm chăm sóc sức khỏe, thể thao và giải trí. Trong lĩnh vực yoga, tính năng ước lượng tư thế có thể được sử dụng để cung cấp phản hồi ngay lập tức cho người tập, giúp họ điều chỉnh tư thế một cách chính xác.

Nghiên cứu của Li và cộng sự [8] đã sử dụng phương pháp từ dưới lên để phân đoạn một học viên yoga trong thời gian thực và ước tính tư thế của nhiều người bằng phương pháp chụp một lần. Họ đã phát triển một CNN đa tác vụ để phát hiện và phân loại chính xác các điểm chính, giúp ước tính tư thế của người tập. Trên bộ dữ liệu gồm 82 tư thế, phương pháp được đề xuất có độ chính xác là 90,1%. Một nghiên cứu khác của Yadav et al. [9] nhằm tạo ra YogNet, một kiến trúc CNN nhẹ được tối ưu hóa cho thiết bị di động, để dự đoán chính xác các tư thế yoga. Nghiên cứu đã sử dụng một tập hợp các mạng thần kinh sâu và đạt được độ chính xác 92,3% trên bộ dữ liệu Yoga-82. Ngoài ra, phương pháp được đề xuất có hiệu quả trong việc điều chỉnh tư thế của học viên, với độ chính xác là 93,7%. Mô hình dựa trên một phần đã được sử dụng, dựa trên cấu trúc cấp điểm chính để đào tạo hoạt động phân đoạn thời gian thực trong nghiên cứu đầu tiên.

Trong hai cuộc điều tra, các vị trí quan trọng được định vị bằng cách sử dụng bản đồ nhiệt và hồi quy bằng cách sử dụng mạng lưới thần kinh sâu để tăng độ chính xác của việc xác định tư thế yoga. Tuy nhiên, có những hạn chế khi đặt ra ước tính, bao gồm sự phụ thuộc vào dữ liệu đầu vào nhất quán, khả năng đơn giản hóa quá mức các tư thế yoga và thiếu cá nhân hóa đối với sự khác biệt về hình dạng, kích thước cơ thể và

tính cơ động. Chỉ dựa vào ước tính tư thế để sửa chữa và ước tính các tư thế yoga có khả năng bỏ qua các sắc thái của nhận thức về hơi thở và tâm trí, đồng thời có thể không đáp ứng được nhu cầu và mục tiêu riêng của từng học viên. Do đó, hướng dẫn được cá nhân hóa từ một giáo viên yoga có trình độ, người có thể cung cấp các điều chỉnh thực hành, tín hiệu bằng lời nói và sửa đổi sẽ bổ sung cho việc sử dụng ước tính tư thế.

Các nhà nghiên cứu đang cố gắng phát triển một hệ thống dựa trên Deep Learning có thể phát hiện chính xác các tư thế yoga và cung cấp cho người dùng phản hồi nhằm nỗ lực thay thế giáo viên. Cuộc điều tra thử nghiệm của King et al. [10] đã xem xét các chiến lược học máy và học sâu để xác định các tư thế yoga. Các mô hình Support Vector Machine, Convolutional Neural Network và Convolutional Neural Network with Long Short-Term Memory đã được thử nghiệm và so sánh kết quả của chúng. Theo nghiên cứu, các mô hình lai CNN-LSTM tạo ra ít phân loại sai nhất.

Microsoft Kinect đã được sử dụng bởi Kadbhane et al. [11] để ghi lại dữ liệu video và xác định 20 khớp cơ thể. Sử dụng dữ liệu được thu thập từ mười vị trí khớp được chọn của con người, cấu trúc tham chiếu cho từng tư thế yoga đã được phát triển. Bằng cách tính toán các góc tạo bởi tất cả các vectơ nối hai vị trí khớp bất kỳ, họ có thể thiết lập sự đồng dạng cosin của các vectơ. Khi dự đoán phân kỳ quá cao, tư thế không chính xác. Tổng độ chính xác đạt được không được đề cập trong bài báo.

S. Haque và cộng sự. [12] đã đề xuất ExNet, một mạng thần kinh tích chập nhiều lớp (CNN). Chồng đẩy, kéo xà, đập xe, uốn cong gân kheo và đi bộ được đưa vào bộ sưu tập 2000 hình ảnh của các cá nhân trong các tư thế yoga. Mô hình đã sử dụng trình tối ưu hóa Adam và kỹ thuật giảm tốc độ học tự động. ExNET có thể xác định tư thế yoga 2D của con người từ bộ dữ liệu với độ chính xác 82,68% sau 50 lần lặp lại. Mô hình có các vấn đề về trang bị quá mức và yêu cầu điều chỉnh siêu tham số lớn hơn. Bộ dữ liệu được sử dụng bởi Agarwal et al. [13] bao gồm 5500 bức ảnh về 10 tư thế yoga riêng biệt. Họ đã đề xuất một phương pháp trong đó bộ xương của người dùng ban đầu được xác định bằng cách sử dụng kỹ thuật ước tính tf-pose. Trong giai đoạn cuối, sáu mô hình học máy khác nhau—Cây quyết định, Rừng ngẫu nhiên, Hồi quy logistic, Naive Bayes, SVM và KNN—đã được sử dụng và đối chiếu. Với tỷ lệ chính xác 99,04%, bộ phân loại Rừng ngẫu nhiên có kết quả tốt nhất.

Anilkumar và cộng sự. [14] đã phát triển một hệ thống sử dụng thư viện Mediapipe để thực hiện phân tích hình học bằng cách sử dụng dữ liệu khung máy ảnh nhằm cung cấp phản hồi cho người dùng về tư thế yoga của họ. Đối với một tư thế yoga cụ thể, hệ thống sẽ tính toán các góc giữa các khớp khác nhau và so sánh chúng với các góc chính xác được ghi lại trong cơ sở dữ liệu. Nếu chênh lệch góc lớn hơn một ngưỡng nhất định, tiện ích sẽ cung cấp phản hồi cho người dùng thông qua văn bản hoặc lời nói. Một hệ thống được tạo bởi Luvizon et al. [15] có thể nhận ra chuyển động của con người ở chiều thứ hai và thứ ba cũng như các tư thế liên quan đến bốn loại hoạt động điển hình khác nhau của con người. Sử dụng đầu vào từ bốn nguồn riêng biệt, hệ thống được đào tạo bằng cách sử dụng mạng thần kinh tích chập đa nhiệm.

Hơn một trăm bản ghi của cả hình thức tập thể dục phù hợp và không phù hợp đã được sử dụng để cung cấp thông tin cho đề xuất [16] của Chen và cộng sự về một hệ thống điều chỉnh tư thế. OpenPose đã được sử dụng để đào tạo mô hình và 18 điểm mốc đã được xác định để nó có thể xác định các hướng cơ thể khác nhau. Chiến lược này đã được đưa vào thử nghiệm với bốn bài tập riêng biệt: nâng cao phía trước, ấn vai đứng, gập bắp tay và nhún vai.

Một huấn luyện viên thể dục thông minh dựa trên phát hiện và ước tính tư thế được phát triển bởi Zou et al. [17] sử dụng thuật toán học sâu. Để so sánh các khớp của con người với các vị trí lý tưởng và cung cấp phản hồi cho người dùng, họ đã sử dụng khung ước tính tư thế nhiều người và AlphaPose. Yadav và cộng sự. [18] cũng sử dụng phương pháp học sâu để phát triển một mô hình thời gian thực giúp xác định chính xác các điểm chính từ các video về sáu tư thế yoga khác nhau. Hệ thống đạt được độ chính xác cao cả về khung và dựa trên các cạnh đặt trước và đã được thử nghiệm trên 12 cá nhân với tỷ lệ thành công là 98,92%. Thoutam và cộng sự. [19] đã sử dụng ước tính tư thế đa mục đích của Keras để trích xuất các tính năng và Perceptron đa lớp để phân loại sáu tư thế yoga dựa trên các góc giữa mười hai điểm chính, đạt được độ chính xác cao 99,58%.

A. Khoảng cách nghiên cứu và động lực

Mặc dù đã có nghiên cứu quan trọng được thực hiện về việc điều chỉnh và ước tính các tư thế yoga bằng cách sử dụng ước tính tư thế với OpenCV, nhưng vẫn còn một số lỗ hổng trong lĩnh vực này. Một trong những lỗ hổng nghiên cứu chính là nhu cầu về các bộ dữ liệu lớn hơn và đa dạng hơn để đào tạo các mô hình học sâu. Hầu hết các bộ dữ liệu hiện tại đều tương đối nhỏ và hạn chế về phạm vi các tư thế yoga mà chúng đề cập. Ngoài ra, thiếu tiêu chuẩn hóa trong các số liệu đánh giá được sử dụng để so sánh các mô hình và kỹ thuật khác nhau, điều này gây khó khăn cho việc so sánh kết quả giữa các nghiên cứu.

Một lỗ hổng nghiên cứu khác là nhu cầu về các thuật toán ước tính tư thế chính xác và mạnh mẽ hơn có thể xử lý các biến thể về ánh sáng, góc máy ảnh và các yếu tố môi trường khác. Nhiều thuật toán ước tính tư thế hiện tại dựa trên các giả định cụ thể về tư thế hoặc môi trường, điều này có thể hạn chế độ chính xác và khả năng ứng dụng của chúng trong các tình huống trong thế giới thực.

Cuối cùng, cần có giao diện thân thiện hơn với người dùng và cơ chế phản hồi để người dùng tương tác với các hệ thống này. Mặc dù nhiều hệ thống hiện tại cung cấp phản hồi dưới dạng văn bản hoặc âm thanh, nhưng cần có giao diện tương tác và thân thiện với người dùng hơn để người dùng có thể nhận được phản hồi và hướng dẫn trong thời gian thực khi tham gia tập yoga.

B. Mục tiêu nghiên cứu

- Thu thập kiến thức về nhiều kỹ thuật tiên tiến để đánh giá lập trường của một người, sau đó chọn một vài khả năng để kiểm tra.

- Xác định những đặc điểm nào là quan trọng nhất để đánh giá dựa trên sự khác biệt về kỹ thuật và loại cơ thể trong môn cử tạ được cho là có nguy cơ chấn thương đáng kể.
- Sản xuất các video thể dục bao gồm các yếu tố kỹ thuật mong muốn cũng như các video không có bất kỳ yếu tố kỹ thuật nào. Sau đó, sử dụng các bộ phim để huấn luyện các thuật toán ước tính tư thế con người, thuật toán này sẽ cung cấp các bộ dữ liệu để thử nghiệm và đánh giá.
- Tạo một ứng dụng web có thể nhận dạng bài tập mà đối tượng đang thực hiện và góc nhìn được sử dụng trong video, cho phép kiểm tra tự động một số yếu tố kỹ thuật.
- Tạo các công thức phổ quát có xác suất phát hiện rủi ro liên quan đến kỹ thuật cao cho người dùng thông thường.
- Xem lại dữ liệu và đối chiếu kết quả của các công cụ ước tính tư thế khác nhau. Kiểm tra khả năng phân biệt phối cảnh và phương pháp quay phim của hệ thống, cũng như khả năng phát hiện một số sai sót nhất định trong kỹ thuật của hệ thống.

III. NGUYÊN LIỆU VÀ PHƯƠNG PHÁP

A. Bộ dữ liệu ước tính tư thế Yoga

Chúng tôi sử dụng bộ dữ liệu trực tuyến có sẵn công khai, YogAI, bao gồm 16 tư thế yoga do 10 người tham gia (5 nam và 5 nữ) thực hiện trong môi trường được kiểm soát với cài đặt ánh sáng và camera nhất quán. Mỗi người tham gia thực hiện các tư thế nhiều lần, tạo ra tổng cộng 1.506 hình ảnh. Các tư thế được chọn để thể hiện một loạt các mức độ khó, bao gồm cả tư thế tĩnh và động. Bộ dữ liệu bao gồm 16 tư thế yoga, bao gồm cả tư thế tĩnh và động. Một số tư thế tĩnh có trong tập dữ liệu là chó săn, cây cối và chiến binh. Một số tư thế năng động có trong bộ dữ liệu là tấm ván và nữ thần.

Đối với mỗi hình ảnh trong bộ dữ liệu YogAI, các chú thích được cung cấp cho các điểm chính trên cơ thể, chẳng hạn như khớp và các bộ phận cơ thể. Các chú thích được tạo bằng công cụ chú thích tùy chỉnh do các nhà nghiên cứu phát triển và được nhiều người chú thích xác minh để đảm bảo độ chính xác. Ngoài các chú thích điểm chính, bộ dữ liệu cũng bao gồm các nhãn tư thế cho mỗi hình ảnh. Chúng tôi chọn 18 điểm chính trên cơ thể, mỗi điểm được tạo thành từ tọa độ x và y của một điểm trên cơ thể. Một từ điển và một mảng 2D được tạo ra bởi điều này. Khi từ điển trả về một số giá trị cho một khóa nhất định, tất cả dữ liệu này cùng với các mức độ tin cậy liên quan sẽ được đưa vào một mảng. Các thành phần vật lý mà các khóa và giá trị của từ điển đại diện là tọa độ của chúng. Sự hiện diện của các điểm cơ thể được quan sát trong từ vựng được xác định dựa trên mức độ tin cậy cao.

Chú thích điểm chính trong tập dữ liệu bao gồm các vị trí khớp (ví dụ: khuỷu tay, đầu gối, mắt cá chân) và các bộ phận cơ thể (ví dụ: đầu, thân, tứ chi). Các chú thích được tạo bằng công cụ chú thích tùy chỉnh do các nhà nghiên cứu phát triển và được nhiều người chú thích xác minh để đảm bảo độ chính xác. Bộ dữ liệu cũng bao gồm nhãn tư thế cho mỗi hình ảnh. Bộ dữ liệu YogAI chứa 1.506 hình ảnh, với trung bình 94 hình ảnh cho mỗi tư thế. Để tạo thuận lợi cho việc đào tạo và đánh giá mô hình, các

tập dữ liệu được chia thành tập huấn luyện (bao gồm 80% ảnh) và tập xác nhận (bao gồm 20% ảnh).

B. Tiền xử lý tập dữ liệu

Chúng tôi xóa các điểm dữ liệu có xác suất chính xác thấp để tránh kết quả dương tính giả. Quá trình lọc liên quan đến việc loại bỏ các điểm được dự đoán là sai dựa trên các phương pháp đánh giá tư thế con người và lọc ra các điểm theo sau có độ biến thiên quá cao. Bước này đóng một vai trò quan trọng trong việc xác định độ chính xác của kết quả và tránh các vấn đề về phương pháp không tồn tại. Để lọc ra những ước tính không chính xác, mỗi phương pháp trong số ba phương pháp ước tính tư thế con người cung cấp một điểm tin cậy cho mỗi lần đoán. Phương pháp đơn giản nhất để loại bỏ để đưa ra các ước tính có độ tin cậy thấp là bỏ qua các chỉ tiết quan trọng. Ngưỡng điểm tin cậy được đặt ở mức 70% để loại bỏ bất kỳ xếp hạng tin cậy nào có xác suất thấp hơn giá trị này. Tuy nhiên, ngưỡng điểm tin cậy lên tới 90% là cần thiết đối với một số tính toán để tính đến nhiều có thể dẫn đến kết quả dương tính giả.

Hệ thống ước tính tư thế con người và thị giác máy tính có thể tạo ra những điểm quan trọng vốn có điểm xác suất cao nhưng có sự khác biệt về khoảng cách quá mức so với khung hình trước đó. Để tránh dữ liệu sai, những điểm này đã được lọc ra. Thông tin về điểm chính được tổ chức thành một từ điển với tọa độ x và y và khả năng nó đứng (z). Chỉ mục của mảng này được sử dụng làm khung và nếu một mục không có, giá trị null sẽ được chèn vào. Điểm tin cậy nằm trong phạm vi [0, 1] và tọa độ x và y được chuẩn hóa để nằm trong phạm vi đó. Ngưỡng điểm tin cậy phải được hạ xuống 60% đối với tất cả các phim được tạo từ góc nhìn chính diện vì điểm tin cậy thấp hơn đối với các điểm quan trọng khi xem chính diện so với phần còn lại của dữ liệu.



Hình 1 Hình ảnh được xử lý trước với Chú thích điểm chính

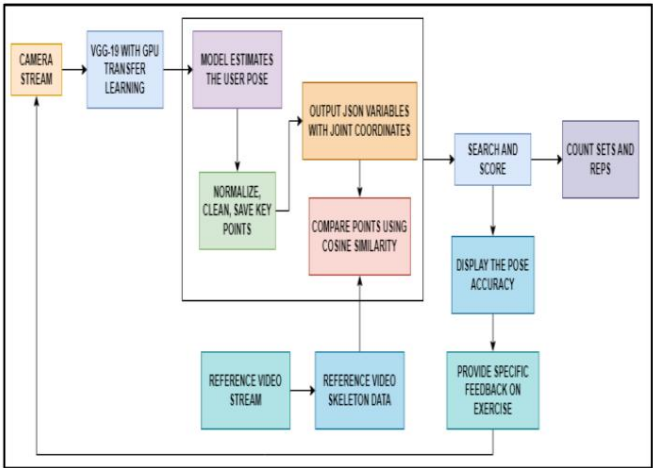
C. Kiến trúc mạng

Kiến trúc mạng được đề xuất cho việc Hiệu chỉnh và Ước tính Tư thế Yoga bằng cách sử dụng Ước tính Tư thế với OpenCV và VGG-19 với GPU transfer learning là sự kết hợp của hai kỹ thuật mạnh mẽ: học sâu và ước tính tư thế Việc sử dụng OpenCV cho hiện tại

đánh giá và VGG-19 để học chuyển động trên nền tảng GPU cung cấp giải pháp chính xác và hiệu quả để điều chỉnh và ước lượng các tư thế yoga. Hình II cho thấy kiến trúc để ước tính và hiệu chỉnh tư thế bằng OpenPose.

Đặt ước tính OpenCV là phần ban đầu của kiến trúc được đề xuất. Một thư viện phổ biến có tên OpenCV được sử dụng cho các ứng dụng thị giác máy tính bao gồm nhận dạng đối tượng, nhận dạng khuôn mặt và ước tính vị trí. Cấu trúc này tận dụng lợi thế của OpenCV để ước tính ngay tư thế 2D của người dùng. Để xác định các khu vực quan trọng của cơ thể con người, bao gồm khớp, tay chân và các thành phần cơ thể khác, OpenCV sử dụng mô hình học sâu được đào tạo trước. Lập trường của người dùng sau đó được ước tính bằng cách sử dụng các điểm thiết yếu này. Độ chính xác của ước tính tư thế là rất quan trọng đối với độ chính xác của toàn bộ hệ thống và OpenCV cung cấp độ chính xác cao trong thời gian thực.

Thành phần thứ hai của kiến trúc được đề xuất là chuyển giao học tập sử dụng VGG-19 trên nền tảng GPU. Bộ dữ liệu ImageNet đã được sử dụng để đào tạo mô hình học sâu được đào tạo trước VGG-19. Việc sử dụng lại một mô hình đã được đào tạo trước đó và tối ưu hóa nó trên một tập dữ liệu mới là những gì mà quá trình học chuyển đổi đòi hỏi để cải thiện độ chính xác. Trong kiến trúc này, học chuyển đổi được sử dụng để tinh chỉnh mô hình VGG 19 trên bộ dữ liệu mới về các tư thế yoga. Việc sử dụng học chuyển đổi trên nền tảng GPU đảm bảo rằng mô hình có thể được đào tạo nhanh hơn và hiệu quả hơn trên nền tảng CPU.

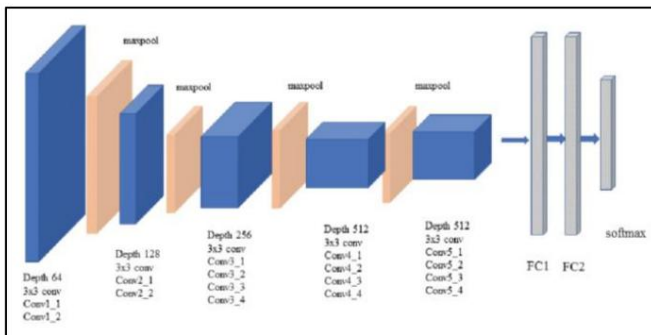


Hình 2 Dòng chảy kiến trúc

Các lớp tích chập tạo nên kiến trúc VGG-19 và mỗi lớp trích xuất thông tin phức tạp dần dần từ ảnh đầu vào. Sau các lớp tích chập, quá trình phân loại cuối cùng được thực hiện bởi ba lớp được ghép nối hoàn toàn. Có tổng cộng 19 lớp trong thiết kế, với 16 lớp đầu tiên đóng vai trò là lớp tích chập và 3 lớp cuối cùng là lớp được kết nối đầy đủ. Các lớp dropout và các lớp tổng hợp tối đa được sử dụng trong thiết kế để giảm kích thước không gian của bản đồ đối tượng địa lý và ngăn chặn trang bị thừa.

Đối với quá trình học chuyển GPU, mẫu VGG-19 được tải lên GPU sau khi nó đã được huấn luyện. Sau khi mô hình đã được đào tạo, trọng số của nó bị khóa và một giá trị mới

tập hợp các lớp được kết nối đầy đủ được thêm vào kiến trúc. Toàn bộ mạng được đào tạo bằng tập dữ liệu mới và trọng số của các lớp bổ sung được khởi tạo ngẫu nhiên. Trọng lượng được đào tạo trước tăng tốc quá trình đào tạo, trong khi GPU tăng tốc tính toán.



Hình 3 Kiến trúc VGG-19 với GPU Transfer Learning

Thành phần cuối cùng của kiến trúc được đề xuất là sự tích hợp của ước tính tư thế và học chuyển giao. Các điểm chính thu được từ mô hình ước tính tư thế OpenCV được sử dụng làm đầu vào cho mô hình VGG-19 để hiệu chỉnh và ước tính các tư thế yoga của người dùng. Mẫu VGG-19 sử dụng các điểm chính làm đặc điểm để phân loại tư thế yoga và cung cấp phản hồi cho người dùng về tính đúng đắn của tư thế. Việc sử dụng mô hình học sâu đảm bảo rằng hệ thống có thể xử lý các biến thể về tư thế, điều kiện ánh sáng và góc camera, cung cấp phản hồi chính xác cho người dùng.

IV. KẾT QUẢ VÀ THẢO LUẬN

Độ chính xác, độ chính xác, độ đặc hiệu và độ nhạy của phân loại của các mạng được đo trong giai đoạn thử nghiệm. Tính toán độ chính xác của phân loại cũng đơn giản như chia số lần đoán chính xác cho tổng số lần dự đoán:

$$\text{độ chính xác} = \frac{\text{+}}{\text{+ + +}}$$

Trong đó, TP = Tích cực Đúng, FP = Tích cực Sai, TN = Tiêu cực Thực sự, FN = Tiêu cực Sai.

Để xác định độ chính xác của một mô hình, một nhà thống kê có thể áp dụng công thức sau đây, so sánh tỷ lệ dự báo chính xác được quan sát và dự kiến.

$$\text{Độ chính xác} = \frac{TP}{TP+FP}$$

Độ nhạy, tỷ lệ các trường hợp tích cực được phân loại chính xác, là chỉ số đo tỷ lệ này. Người ta có thể biểu thị độ nhạy theo thuật ngữ toán học như

$$TPR = \frac{TP}{TP + FN}$$

Chúng tôi cũng xây dựng các giá trị âm tính thực, cho biết tỷ lệ các trường hợp âm tính giả được phát hiện chính xác dựa trên loại của chúng, để đánh giá tính đặc hiệu:

$$TNR = \frac{TN}{TN+FP}$$

Trong nghiên cứu, dữ liệu đào tạo được gấp 10 lần hoặc chia đều để tối đa hóa hiệu quả mà theo đó có thể học được ánh xạ giữa đầu vào và đầu ra. Các siêu tham số của mô hình đã được tối ưu hóa bằng cách sử dụng các nếp gấp này và hiệu suất của hệ thống học sâu được đánh giá bằng nhiều biện pháp khác nhau. Điểm số cho MobileNet Tinh chỉnh và VGG-19 với Kiến trúc GPU Transfer Learning đã được tính hơn 10 lần về độ chính xác, độ đặc hiệu, độ nhạy và độ chính xác. Cân nhắc mức giảm trọng lượng là 0,00001, quy trình đào tạo đã sử dụng mười tỷ lệ học tập riêng biệt: 0,01 cho 100 kỷ nguyên đầu tiên, 0,001 cho 100 kỷ nguyên tiếp theo, v.v.

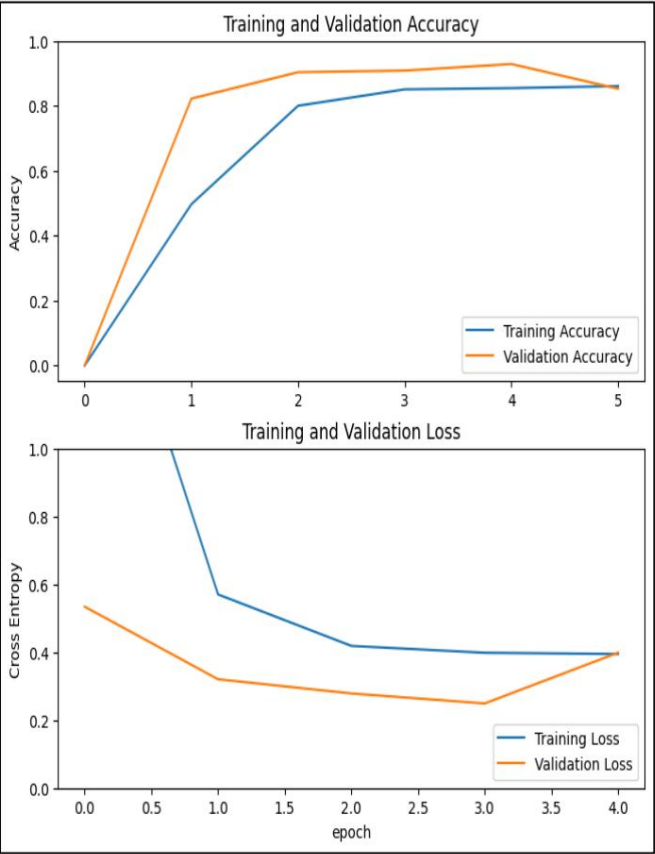
Đào tạo chiếm 70% thời gian trong nghiên cứu, với thử nghiệm chiếm 10% và xác nhận chiếm 20% nỗ lực. Tập dữ liệu được thu thập từ nhiều lĩnh vực khác nhau và được phân chia thành các tập kiểm tra và đào tạo dựa trên chủ đề để ngăn thông tin từ một nhóm ảnh hưởng đến nhóm khác. Độ chính xác ổn định giữa 800 và 1000 lần lặp lại, Do đó, Xác thực chéo 10 lần đã được thực hiện để đánh giá hiệu quả của bộ cục được đề xuất. với 100 lần lặp lại được tính cho đào tạo trong mỗi lần.

Để đảm bảo kiểm tra kỹ lưỡng hệ thống, một số quy trình có thể đã được sửa đổi. Tuy nhiên, bất chấp các phương pháp khác nhau được sử dụng, các phát hiện của nghiên cứu cho thấy một giải pháp thỏa đáng. Tuy nhiên, bộ dữ liệu được sử dụng trong nghiên cứu có thể được hưởng lợi từ việc bao gồm nhiều cá nhân và video hơn để giải quyết vấn đề nghiên cứu một cách thỏa đáng. Để phần mềm này có thể bán được trên thị trường, nó cần được người tiêu dùng chấp nhận. Các điểm chính trên video 2 chiều cung cấp hướng dẫn quan trọng về việc liệu người dùng có đang sử dụng hình thức phù hợp trong suốt quá trình tập yoga của họ hay không. Tuy nhiên, điều cần thiết là phải biết cơ nào sẽ tham gia trong một bài tập cử tạ. Mặc dù hình thức phù hợp có thể làm tăng khả năng sử dụng đúng cơ, nhưng thiết bị phức tạp khác hoặc đầu vào của người dùng có thể cần thiết để xác định điều này.

Các phát hiện của nghiên cứu cho thấy rằng Ước tính tư thế người 2D có thể đưa ra phản hồi trực tiếp về phương pháp tập tạ ở những người khỏe mạnh. Tuy nhiên, cần nghiên cứu thêm để xác định xem những kết quả này có thể được dịch sang các góc nhìn bên và để cải thiện khả năng phát hiện kỹ thuật hay không. Việc triển khai thành công các chế độ xem bên của Pose Trainer gợi ý rằng độ cong thời gian động thích hợp hơn cho các chi tiết cụ thể của kỹ thuật như vậy. Việc nhận biết các tác vụ yêu cầu nhiều vòng quay có thể khó khăn, nhưng thuật toán vẫn cung cấp thông tin chi tiết hữu ích. Cách tiếp cận hiện tại tạo ra kết quả tương tự như Pose Trainer trong khi thu được từ tập dữ liệu lớn hơn và đưa ra phản hồi chi tiết hơn về các bài tập khó hơn. Nghiên cứu này mở rộng dựa trên những nỗ lực trước đó với Pose Trainer bằng cách chứng minh các phản quan trọng của Ước tính tư thế người 2D có thể được tích hợp với các phương pháp cơ bản hơn để đưa ra phản hồi về hình thức cử tạ phù hợp.

Bảng 1 Số liệu hiệu suất cho VGG-19 với Chuyển giao

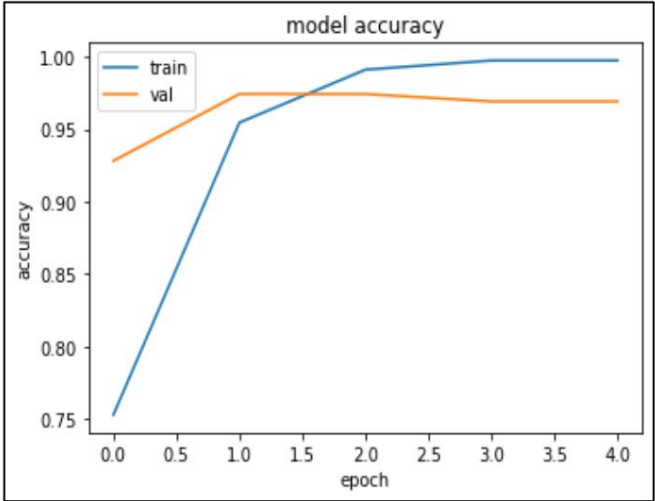
nếp gấp	Số liệu hiệu suất			
	Nó chỉ định	Ảnh áy cảm thấy nó	chính xác	tôi cần
	Và	Và	Và	N
Gấp-I	98,78	98,95	98,61	98,86
Gấp-II	98,41	98,45	98,52	98,22
Gấp-III	98,99	98,96	98,39	98,54
Gấp-IV	98,36	98,81	98,65	98,43
Gấp-V	98,74	98.11	98,73	98,26
Gấp-VI	98,90	98.86	98,84	98,72
Gấp-VII	98,22	97.68	97,68	96,22
Gấp-VIII	97,99	97.59	97,93	97,23
Gấp-IX	98,70	97.95	98,65	96,56
Gấp-X	97,76	98.16	98,46	96,74
chồng chéo d dữ liệu	NULL NULL	NULL NULL		
Trung bình	98.11	97,42	98,37	97,75



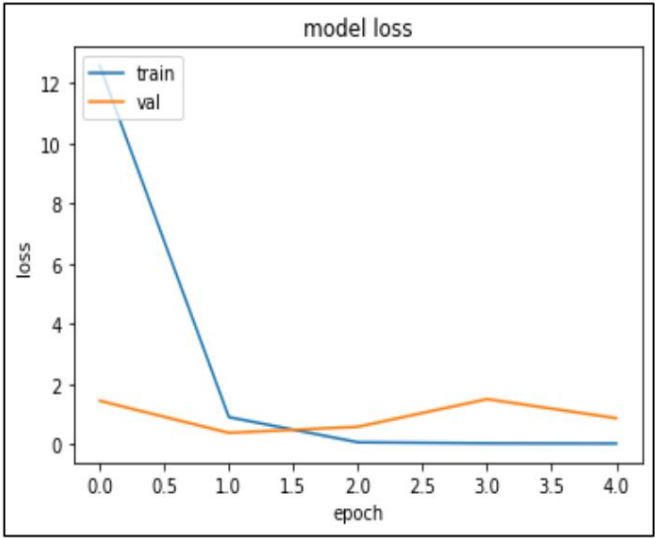
Hình 4 Độ chính xác đào tạo và đường cong tổn thất cho VGG-19 với Học chuyển GPU

Rõ ràng là việc sử dụng Ước tính tư thế người 2D và các phương pháp học máy để nghiên cứu về tập tậ có tiềm năng trở thành một công cụ hữu ích để giảm nguy cơ chấn thương và nâng cao hiệu quả tập luyện tổng thể. Những phát hiện của nghiên cứu này cho thấy rằng một bộ dữ liệu lớn hơn và những cải tiến đối với phát hiện kỹ thuật có thể dẫn đến kết quả chính xác hơn, đặc biệt là phát hiện chế độ xem bên. Ngoài ra, có thể là sự biến dạng động của hệ thống có thể được sử dụng tốt hơn bằng cách tiếp cận khai thác nó. Lưu ý rằng dự đoán góc là một phần quan trọng trong việc đánh giá phim đào tạo và dự đoán sai có thể dẫn đến kết quả xấu.

kết quả là điều cần thiết. Tuy nhiên, phương pháp được sử dụng trong nghiên cứu này đã dự đoán thành công mọi điểm thuận lợi bằng cách tính toán hướng của vectơ giữa hai vai. Với tập dữ liệu lớn hơn và phân phối tương đương các chế độ xem chính diện và ngang, phương pháp này có thể mang lại kết quả tương tự.



Hình 5 Đường cong đào tạo chính xác cho MobileNet tinh chỉnh Ngành kiến trúc



Hình 6 Đường cong tổn thất đào tạo cho MobileNet tinh chỉnh Ngành kiến trúc

Điều quan trọng là phải ghi nhớ những hạn chế này và các nguồn lỗi tiềm ẩn khi sử dụng máy dò góc. Có thể cần thử nghiệm và cải tiến thêm thuật toán để tăng độ chính xác và khả năng phát hiện phạm vi góc rộng hơn. Về hệ thống đánh giá kỹ thuật, hứa hẹn rằng phương pháp học máy mà Pose Trainer sử dụng có thể xác định chính xác dạng bài tập đúng và sai cho một số bài tập. Tuy nhiên, cần lưu ý rằng nghiên cứu dựa trên một tập dữ liệu nhỏ bao gồm hầu hết những người khỏe mạnh. Nghiên cứu và thử nghiệm bổ sung có thể cần thiết để xác định hiệu quả của hệ thống đối với nhiều loại bài tập và cá nhân với các mức độ thể chất và sức khỏe khác nhau.



Hình 7 Ước tính tư thế Yoga bằng VGG-19 với GPU học chuyển tiếp

TRONG

PHẦN KẾT LUẬN

Tóm lại, Hiệu chỉnh và Ước tính Tư thế Yoga bằng cách sử dụng Ước tính Tư thế với OpenCV là một công nghệ đầy hứa hẹn để giúp các cá nhân cải thiện việc luyện tập yoga của họ. Nghiên cứu này đã chứng minh tính hiệu quả của việc sử dụng Ước tính tư thế người 2D để phát hiện và điều chỉnh các tư thế yoga bằng cách cung cấp phản hồi thời gian thực cho người dùng. Pose Trainer được phát triển trong nghiên cứu này đã cho thấy thành công trong việc xác định cách thực hiện đúng và sai các tư thế yoga khác nhau, đồng thời cung cấp phản hồi chính xác cho người dùng.

Những phát hiện của nghiên cứu này cũng cho thấy rằng Pose Trainer có thể hữu ích trong việc giảm nguy cơ chấn thương liên quan đến các bài tập tạ. Bằng cách cung cấp phản hồi về kỹ thuật phù hợp và sự tham gia của cơ bắp, các cá nhân có thể thực hiện các bài tập cử tạ với hình thức phù hợp và giảm thiểu nguy cơ chấn thương.

Tuy nhiên, vẫn còn nhiều chỗ để cải thiện trong công nghệ này. Tập dữ liệu được sử dụng trong nghiên cứu này có thể được hưởng lợi từ việc lớn hơn và đa dạng hơn, đồng thời cần có nhiều nghiên cứu hơn để xác định xem liệu kết quả từ tính năng phát hiện chế độ xem trước có thể được dịch sang tính năng phát hiện chế độ xem bên hay không. Ngoài ra, hệ thống phát hiện góc được sử dụng trong nghiên cứu này có những hạn chế và có thể không chính xác trong việc phát hiện các góc không trực diện hoặc xiên.

Nhìn chung, nghiên cứu này đã đóng góp vào lĩnh vực ước tính tư thế và thị giác máy tính đang phát triển và có khả năng cung cấp hỗ trợ có giá trị cho những cá nhân muốn cải thiện việc luyện tập yoga hoặc thực hiện các bài tập cử tạ một cách an toàn. Với việc tiếp tục nghiên cứu và phát triển, công nghệ này có thể trở thành một công cụ có giá trị trong ngành y tế và thể dục.

NGƯỜI GIỚI THIỆU

- [1]. Hồi giáo, Muhammad Usama & Mahmud, Hasan & Ashraf, Faisal Bin & Hossain, Iqbal & Hasan, Md Kamrul. (2017). Nhận dạng tư thế yoga bằng cách phát hiện các điểm khớp của con người trong thời gian thực bằng microsoft kinect. 668-673. 10.1109/R10-HTC.2017.8289047.
- [2]. H. Tang, Q. Wang và H. Chen, "Research on 3D Human Pose Estimation Using RGBD Camera," 2019 IEEE 9th International Conference on Electronics Information and Emergency Communication (ICEIEC), doi: 2019, 538-541, 10.1109/ICEIEC.2019.8784591. tr.
- [3]. Triết Cao et al. "OpenPose: Ước tính tư thế 2D nhiều người trong thời gian thực bằng cách sử dụng các trường sở thích một phần". Trong: IEEE Transactions on Pattern Analysis and Machine Intelligence 43.1 (2021), trang 172-186. doi: 10.1109/TPAMI.2019.2929257.
- [4]. Santiago Riofrio và cộng sự. "Nhận dạng cử chỉ bằng cách sử dụng Dynamic Time Warping và Kinect: Một cách tiếp cận thực tế". Trong: Tháng 11 năm 2017, trang 302-308. doi: 10.1109/INCISCOS.2017.36.
- [5]. TL Munea, YZ Jembre, HT Weldegebriel, L. Chen, C. Huang và C. Yang, "Tiến trình ước tính tư thế người: Khảo sát và phân loại các mô hình được áp dụng trong ước tính tư thế người 2D," trong IEEE Access, tập. 8, trang 133330-133348, 2020, doi: 10.1109/ACCESS.2020.3010248.
- [6]. Jinbao Wang, Shujie Tan, Xiantong Zhen, Shuo Xu, Feng Zheng, Zhenyu He, Ling Shao, Ước tính tư thế con người 3D sâu: Đánh giá, Thị giác máy tính và Hình ảnh, Tập 210, 2021, 103225, ISSN 1077- 3142, <https://doi.org/10.1016/j.cviu.2021.103225>.
- [7]. Valcarce-Torrente M, Javaloyes V, Gallardo L, García-Fernández J, Planas-Anzano A. Ảnh hưởng của các ứng dụng thể dục đối với thói quen thể thao, sự hài lòng và ý định ở lại của người dùng trung tâm thể hình: Một nghiên cứu thực nghiệm. Int J Environ Res Y tế công cộng. 2021 ngày 2 tháng 10;18(19):10393. doi: 10.3390/ijerph181910393. PMID: 34639692; PMCID: PMC8507994.
- [8]. Li, Miaopeng & Zhou, Zimeng & Li, Jie & Liu, Xinguo. (2018). Ước tính tư thế từ dưới lên của nhiều người với ràng buộc hộp giới hạn. 115-120. 10.1109/ICPR.2018.8546194.
- [9]. Yadav, Santosh & Agarwal, Aayush & Kumar, Ashish & Tiwari, Kamlesh & Pandey, Hari & Ali Akbar, Shaik. (2022). YogNet: Mạng hai luồng để nhận dạng hành động yoga nhiều người theo thời gian thực và điều chỉnh tư thế. Hệ thống dựa trên tri thức. 250. 109097. 10.1016/j.knosys.2022.109097.
- [10]. S. Kinger, A. Desai, S. Patil, H. Sinalkar và N. Deore, "Phân loại tư thế yoga dựa trên học sâu", Hội nghị quốc tế năm 2022 về học máy, dữ liệu lớn, đám mây và điện toán song song (COM-IT-CON), Faridabad, Ấn Độ, 2022, doi: 10.1109/COM-IT 682-691, tr.

CON54601.2022.9850693.

- [11]. Kadbhane, S., Datir, K., Jagdale, T., Dhongade, S., & Jagtap, G. (2021). Nhận dạng tư thế Yoga. Tạp chí Quốc tế về Nghiên cứu Cao cấp về Kỹ thuật Máy tính và Truyền thông, 10(1), 143-147. doi:10.17148/IJARCCCE.2021.10128
- [12]. Haque, S., Rabby, AS, Laboni, MA, Neehal, N., & Hossain, SA (2019). ExNET: Mạng thần kinh sâu để phát hiện tư thế tập thể dục. Truyền thông trong Khoa học Máy tính và Thông tin Xu hướng Gần đây trong Xử lý Hình ảnh và Nhận dạng Mẫu, 186-193. doi:10.1007/978-981-13-9181-1_17
- [13]. Agrawal, Yash & Shah, Yash & Sharma, Abhishek. (2020). Triển khai Kỹ thuật học máy để xác định các tư thế Yoga. 40-43. 10.1109/CSNT48778.2020.9115758.
- [14]. Anilkumar, Ardra & KT, Athulya & Sajan, Sarath & KA, Sreeja. (2021). Đặt ra Hệ thống giám sát Yoga ước tính. Tạp chí điện tử SSRN. 10.2139/ssrn.3882498.
- [15]. Luvizon, Diogo C., David Picard và Hedi Tabia. "Ước tính tư thế 2d/3d và nhận dạng hành động bằng cách sử dụng học sâu đa tác vụ." Kỷ yếu hội nghị IEEE về thị giác máy tính và nhận dạng mẫu. 2018
- [16]. Chen, Steven & Yang, Richard. (2018). Pose Trainer: Chính sửa tư thế tập thể dục bằng cách sử dụng Ước tính tư thế.
- [17]. Zou, J., Li, B., Wang, L., Li, Y., Li, X., Lei, R., & Sun, S. (2019). Hệ thống Huấn luyện viên Thể hình Thông minh Dựa trên Ước tính Tư thế Con người. Xử lý thông tin và tín hiệu, Mạng và máy tính, 593- 599. doi:10.1007/978-981-13-7123-3_69
- [18]. Yadav, Santosh & Singh, Amitojdeep & Gupta, Abhishek & Raheja, Jagdish. (2019). Nhận dạng Yoga theo thời gian thực bằng cách sử dụng học sâu. Điện toán thần kinh và ứng dụng. 31. <https://link.springer.com/article/10.1007/s00521-019-04232-7>. 10.1007/s00521-019-04232-7.
- [19]. Vivek Anand Thoutam, Anugrah Srivastava, Tapas Badal, Vipul Kumar Mishra, GR Sinha, Aditi Sakalle, Harshit Bhardwaj, Manish Raj, "Ước tính tư thế Yoga và tạo phản hồi bằng cách sử dụng học sâu", Trí tuệ tính toán và khoa học thần kinh, tập. 2022, Article ID 4311350, 12 trang, 2022. <https://doi.org/10.1155/2022/4311350>
- [20]. SK Yadav, A. Singh, A. Gupta và JL Raheja, "Nhận dạng Yoga trong thời gian thực bằng cách sử dụng học sâu," Điện toán thần kinh & Ứng dụng, tập. 31, không. 12, trang 9349-9361, 2019.