



Phát hiện và chỉnh sửa tư thế Yoga bằng Posenet và KNN

Varsha Bhosale¹, Pranjal Nandeshwar², Abhishek Bale³, Janmesh Sankhe⁴

¹Phó giáo sư, ^{2,3,4}Sinh viên mới tốt nghiệp

^{1,2,3,4}Khoa Công nghệ Thông tin, Học viện Công nghệ Vidyalandkar, Mumbai, Ấn Độ.

Tóm tắt - Mục tiêu cơ bản của việc phát hiện và hiệu chỉnh tư thế Yoga là cung cấp các tư thế yoga chuẩn và đúng bằng cách sử dụng thị giác máy tính. Nếu tư thế yoga không được thực hiện đúng cách, nó có thể dẫn đến chấn thương nghiêm trọng và các vấn đề lâu dài. Phân tích các tư thế của con người để phát hiện và điều chỉnh các tư thế yoga có thể mang lại lợi ích cho con người khi sống một cuộc sống lành mạnh hơn trong môi trường gia đình của họ. Dự án này tập trung vào việc khám phá các cách tiếp cận khác nhau để phân loại tư thế yoga, vì vậy chúng tôi đang sử dụng trình phân loại PoseNet và KNN. Sử dụng các thuật toán học sâu như vậy, một cá nhân có thể có được cách thức/phương pháp chính xác/lý tưởng để thực hiện tư thế yoga cụ thể mà họ đang cố gắng thực hiện. Sử dụng các kỹ thuật thị giác máy tính và Open Pose (một thư viện mã nguồn mở), ước tính tư thế con người được sử dụng để ước tính tư thế Yoga của một cá nhân. Hệ thống được đề xuất nhận ra sự khác biệt giữa vị trí thực tế và mục tiêu, đồng thời sửa lỗi cho người dùng với độ chính xác cao bằng cách cung cấp đầu ra trực quan theo thời gian thực và các hướng dẫn cần thiết để sửa tư thế đã xác định.

Từ khóa: KNN, Open Pose, Posenet, Pose Estimation, Crucial Angles.

1. GIỚI THIỆU

Yoga, một môn tập luyện hàng thế kỷ có nguồn gốc từ Ấn Độ nhưng nổi tiếng toàn cầu vì vô số lợi ích về tinh thần, thể chất và tinh thần, là một loại bài tập với các tư thế phức tạp. Vấn đề với yoga là, giống như bất kỳ bài tập nào khác, điều quan trọng là phải thực hành nó đúng cách vì bất kỳ tư thế sai nào trong buổi tập yoga đều có thể không hiệu quả và có khả năng gây bất tiện. Điều này đòi hỏi sự có mặt của một huấn luyện viên để giám sát cuộc họp và điều chỉnh lập trường của cá nhân. Vì không phải mọi khách hàng đều tiếp cận hoặc có quyền tiếp cận với huấn luyện viên nên một ứng dụng dựa trên lý luận được vi tính hóa có thể được sử dụng để phát hiện các tư thế yoga và cung cấp phản hồi tùy chỉnh để giúp mọi người cải thiện cấu trúc của họ. Do đó, chúng tôi đã đề xuất một ứng dụng dựa trên trí tuệ nhân tạo có thể nhận dạng các tư thế yoga và phản hồi bằng phản hồi tùy chỉnh để giúp người dùng cải thiện tư thế của họ. Mục đích của nghiên cứu này là để hiểu rõ hơn về nhiều kỹ thuật được sử dụng để phân loại các tư thế yoga, bao gồm ước tính tư thế, phân loại KNN [1], Posenet [2] và mô hình phân loại KNN [3]. Thiếu thể chất làm tăng nguy cơ mắc các tình trạng sức khỏe bất lợi bao gồm bệnh tim mạch vành, huyết áp cao, đột quỵ, hội chứng chuyển hóa, tiểu đường tuýp 2 dẫn đến giảm tuổi thọ của con người. Trong công việc này, ước tính tư thế để điều chỉnh các tư thế yoga được thực hiện. Nhiều hệ thống phát hiện tư thế là mã nguồn mở

yêu cầu thiết lập hệ thống và camera. Đây là mô hình yoga cơ bản cho phép mọi người thực hiện các tư thế yoga khác nhau và điều chỉnh tư thế yoga của họ bằng cách nhìn vào hình ảnh phản chiếu.

2. KHẢO SÁT TÀI LIỆU

Học sâu là một lĩnh vực nghiên cứu đầy hứa hẹn cho phép chúng ta kiểm tra một lượng lớn dữ liệu theo cách có thể mở rộng. Học sâu, không giống như các mô hình học máy truyền thống, không yêu cầu khai thác tính năng hoặc kỹ thuật vì nó hiểu các mẫu phức tạp trong dữ liệu và tự trích xuất các tính năng. Công việc trước đó nhấn mạnh vào cái nhìn trực quan của các cấu trúc, mô tả bản chất sáng tạo của con người như một sự sắp xếp vững chắc. Con người với bộ xương kết nối các bộ phận cơ thể của họ được sử dụng để thu nhận kiến thức vào cơ thể con người. Để giả định rằng mỗi bộ phận cơ thể độc lập với các bộ phận khác, mô hình cơ thể động học đặt trước thường được sử dụng.

Các tiện ích mở rộng mở rộng gần đây đã được tạo ra, chẳng hạn như các mô hình xuất hiện hỗn hợp, phân cấp, đa phương thức và mạnh mẽ, chẳng hạn như dự đoán giảm / tuần tự. Mỗi chi được thể hiện bằng một bố cục bao gồm vị trí, hướng, sắp xếp hình dạng và mô hình bố trí. Patil et al. [4] đã đề xuất một kỹ thuật để xác định sự khác biệt về tư thế yoga giữa chuyên gia và người tập bằng cách sử dụng thông tin đường viền hình ảnh và tăng tốc các tính năng mạnh mẽ (SURF). Tuy nhiên, việc mô tả và so sánh các tư thế gần như hoàn toàn dựa trên dữ liệu đường viền là không đủ.

Phương pháp nhận dạng Yoga sử dụng máy ảnh RGB điển hình đã được đề xuất bởi Santosh Kumar Yadav et al. [5]. Họ đã sử dụng một webcam độ nét cao để thu thập dữ liệu. Sử dụng OpenPose, họ đã quét người dùng và phát hiện ra các điểm chính. Lớp CNN phân tán theo thời gian được sử dụng để nhận dạng các mẫu giữa các điểm chính trong một khung hình và LSTM được sử dụng để ghi nhớ các mẫu được phát hiện trong các khung hình gần đây. Phương pháp của họ đã loại bỏ yêu cầu về Kinect hoặc bất kỳ công nghệ cụ thể nào khác để xác định các tư thế Yoga.

Quách và cộng sự. [6] đã trình bày một hệ thống thông minh dựa trên hình ảnh và văn bản dành cho việc luyện tập yoga bao gồm các đơn vị đo lường quán tính (IMU) và máy kéo. Tuy nhiên, điều này có thể gây bất tiện cho người dùng, cũng như phá vỡ tư thế yoga tự nhiên và họ đã không cân nhắc tư thế của người tập.

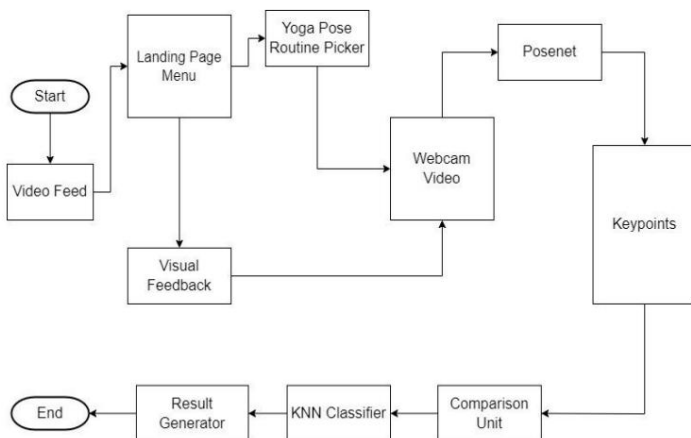
[7] đã mô tả một kỹ thuật khác để chỉnh sửa tư thế yoga dựa trên Kinect có tính đến ba tư thế yoga: chiến binh III, tư thế chó úp mặt và tư thế cái cây. Tuy nhiên, điểm chính xác của chúng chỉ là 82,84%, không mấy ấn tượng. Các công nghệ dựa trên học sâu hiện đã thay thế quy trình khung xương cũ.

Shruti Kothari [8] đã mô tả một phương pháp tính toán để phân loại các tư thế Yoga từ các bức ảnh sử dụng phương pháp học sâu, cụ thể là CNN. Đối với mô hình phân loại, họ đã sử dụng bộ dữ liệu gồm 1000 ảnh được chia thành sáu loại. Công việc này đã được hoàn thành với độ chính xác gần 85%.

Chaudhari và cộng sự. [9] đã xây dựng một hệ thống cung cấp phản hồi rõ ràng cho người tập để họ có thể thực hành đúng các tư thế yoga bằng cách sử dụng kiến thức chuyên môn về năm tư thế yoga. Một mô hình CNN đã được sử dụng để xác định các tư thế yoga và một mô hình bản địa hóa khớp người được sử dụng để tìm ra các khiếm khuyết trong tư thế.

3. HỆ THỐNG ĐỀ XUẤT

Bộ dữ liệu, bao gồm một bộ sưu tập gần 400 Hình ảnh riêng biệt của mỗi Tư thế Yoga. Những Hình ảnh này được tối ưu hóa thành các pixel và kích thước tương tự, sau đó được trộn thành một video. Mô hình Đề xuất sẽ được đào tạo bằng cách sử dụng video đầu ra bao gồm hình ảnh của các tư thế khác nhau trong đó bộ phân loại KNN được sử dụng để phát hiện tư thế bằng cách phát hiện các điểm chính của các chi của con người và mô hình được lưu lại. Chúng tôi đang sử dụng webcam làm nguồn cấp dữ liệu đầu vào. Sử dụng webcam, chúng tôi có thể nhận được Tư thế Yoga của người dùng theo thời gian thực và trình phân loại KNN sẽ xác định các điểm quan trọng và sẽ phân loại nó thành một trong các Tư thế Yoga, sau đó chúng tôi có thể xem video hướng dẫn thực hiện Tư thế Yoga được phát hiện. Chúng ta có thể xem sơ đồ kiến trúc bên dưới.



Hình 1: Sơ đồ kiến trúc

3.1 Trình phân loại KNN

K-Nearest Neighbors là một thuật toán Machine Learning lưu trữ tất cả các trường hợp khác nhau và phân loại dựa trên sự tương đồng. Nó chủ yếu được sử dụng trong kỹ thuật Nhận dạng mẫu. Thuật toán KNN dự đoán giá trị của các điểm dữ liệu mới dựa trên "tính năng tương tự", ngụ ý rằng điểm dữ liệu mới sẽ được gán một giá trị tùy thuộc vào mức độ khớp của nó với các điểm trong tập huấn luyện. Hoạt động của nó được thể hiện bằng cách sử dụng các thủ tục dưới đây:

1. Cần có một bộ dữ liệu để thực hiện bất kỳ thuật toán nào. Do đó, dữ liệu đào tạo và kiểm tra phải được tải trong bước đầu tiên của KNN.
2. Giá trị của K, tức là các điểm dữ liệu gần nó nhất. K có thể là bất kỳ số nguyên nào được yêu cầu chọn.
3. Thực hiện các thao tác sau đối với từng điểm trong dữ liệu thử nghiệm:
 - i) Tính khoảng cách giữa mỗi hàng của dữ liệu huấn luyện và dữ liệu kiểm tra bằng một trong các phương pháp sau: khoảng cách Euclidean, Manhattan hoặc Hamming. Phương pháp Euclidean là phương pháp được sử dụng rộng rãi nhất để tính khoảng cách.
 - ii) Sắp xếp chúng theo thứ tự tăng dần tùy thuộc vào giá trị khoảng cách.
 - iii) Sau đó, nó chọn K hàng trên cùng từ mảng đã sắp xếp.
 - iv) Danh mục thường xuyên nhất của các hàng này sẽ được sử dụng để chỉ định danh mục cho điểm kiểm tra.

3.2 Bộ dữ liệu

Bộ dữ liệu được sử dụng cho việc này được thu thập từ Kaggle [11] và các nguồn mở khác. Nó bao gồm hình ảnh của 5 tư thế yoga khác nhau là - Utkata Konasana (nữ thần), Tadasana (núi), Phalakasana (tám ván), Vrksasana (cây), Virabhadrasana III (chiến binh 3). Tổng cộng là không. trong số 1578 Hình ảnh được thu thập.

3.3 TỬ

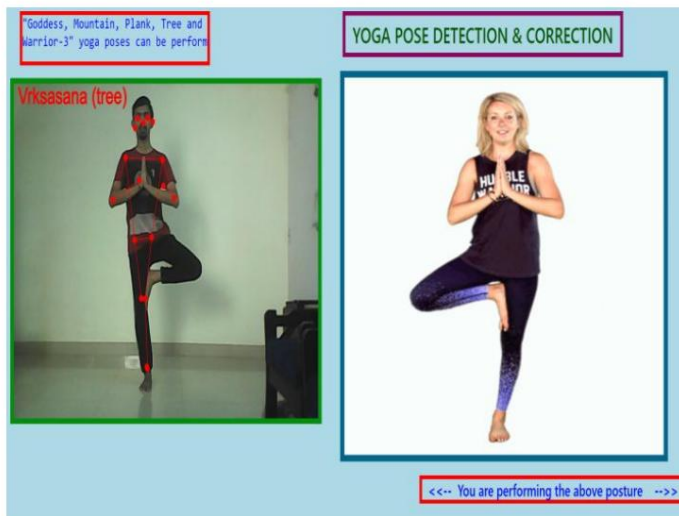
Đó là khung học sâu được sử dụng để xác định tư thế của con người trong chuỗi hình ảnh và video bằng cách xác định các điểm quan trọng trên cơ thể con người và Huấn luyện Người mẫu dựa trên các điểm chính này.

3.4 Mô hình đào tạo

Một video được tối ưu hóa với tất cả các tư thế yoga riêng biệt được tạo bằng các thư viện python như TensorFlow - Keras, OpenPose, NumPy. Giờ đây, Deep learning framework sẽ trích xuất các điểm chính từ chuỗi video để chúng tôi có thể phân loại và đào tạo mô hình bằng giao diện người dùng Frontend. Độ chính xác đào tạo: 93,56%

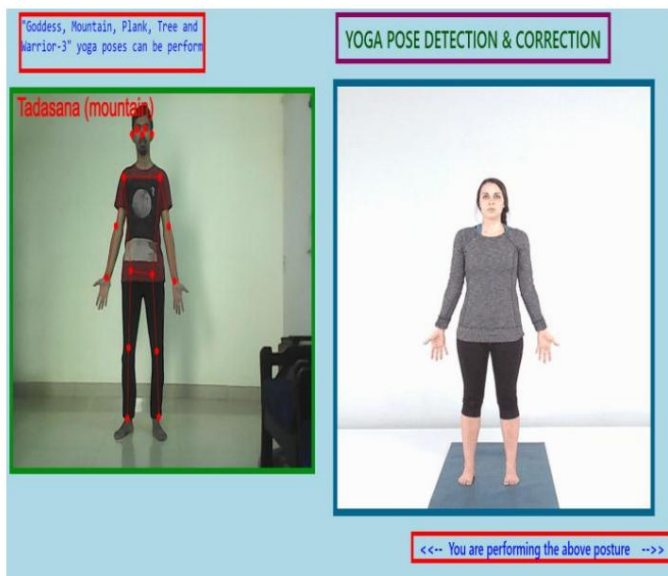
3.5 Hiệu suất mô hình

Sử dụng webcam, Người mẫu sẽ trích xuất các điểm chính từ Tư thế Yoga do người dùng thực hiện. Tùy thuộc vào nguồn cấp dữ liệu, nó sẽ phân loại nó thành một trong các Tư thế Yoga được đào tạo và sau đó chúng ta có thể xem video Người hướng dẫn Yoga thực hiện chính xác tư thế được phát hiện để người dùng có thể tìm hiểu và sửa tư thế cho phù hợp. Kết quả rất hứa hẹn với tỷ lệ chính xác là 98,51%.



Hình 2: Tư thế Vrksasana (cái cây)

Trong hình trên, khi người dùng đang thực hiện Vrksasana (cây), các điểm chính của người dùng như mắt, tai, mũi, tay chân được so sánh với người mẫu đã được huấn luyện. Asana yoga có độ tin cậy cao nhất sẽ được hiển thị trên màn hình. Khi người dùng đang thực hiện Vrksasana (cây), tư thế đứng của người đó cùng với các bước sẽ được hiển thị để có thể hỗ trợ người dùng thực hiện tư thế yoga.



Hình 3: Tư thế Tadasana (Ngon núi)

Trong hình trên, chúng ta có thể thấy người dùng đang thực hiện Tư thế Tadasana (Ngon núi), các điểm mẫu chốt của các chi được phát hiện và mô hình được đào tạo bởi KNN classifier sẽ phân loại nó thành một trong các tư thế đã lưu, vì chúng ta có thể thấy một người hướng dẫn đang thực hiện tư thế tương tự như vậy người dùng đó có thể làm theo các bước và sửa tư thế của mình.



Hình 4: Tư thế Phalakasana (Plank)

Người dùng đang thực hiện Tư thế Phalakasana (Plank), ở phía bên phải, chúng ta có thể thấy một video hướng dẫn thực hiện tư thế tương tự được hiển thị.

4. KẾT LUẬN

Trong vài năm qua, rất nhiều nghiên cứu đã đi vào ước tính vị trí của con người. Ước tính tư thế con người khác với các thử thách thị giác máy tính khác ở chỗ nó phải xác định vị trí và lắp ráp các bộ phận cơ thể con người dựa trên cấu trúc cơ thể con người hiện có. Việc sử dụng ước tính tư thế yoga trong thể dục có thể giúp mọi người tránh chấn thương và tăng khả năng tính hiệu quả. Các thuật toán phân loại hình ảnh hiện đại như KNN được đưa vào thử nghiệm trong hệ thống này cùng với Posenet. Kết quả khá ấn tượng với độ chính xác 98,51%. Chuyển động của các Asana yoga có thể được phân tích bằng phân tích video và hình ảnh để kiểm tra tính chính xác của chúng. Các thiết kế mô hình như Posenet, M15 và KNN Classifier phù hợp để phân tích dựa trên video.

5. CÔNG VIỆC TƯƠNG LAI

Hiện tại, mô hình được đề xuất phân loại các asana yoga thành 5 loại. Vì có rất nhiều asana yoga nên việc phát triển một mô hình ước tính tư thế phù hợp với tất cả chúng là một công việc tẻ nhạt. Nhiều tư thế yoga do các cá nhân thực hiện ở cả trong nhà và ngoài trời có thể được thêm vào tập dữ liệu. Việc hiển thị mô hình được xác định bởi bản chất của đánh giá hiện tại của Open Pose, có thể hoạt động tốt hoặc không trong các tình huống khi có sự chồng chéo giữa các người hoặc giữa các bộ phận cơ thể. Khung này có thể được triển khai trên thiết bị di động để tự chuẩn bị và dự đoán liên tục. Đối với các ứng dụng hợp lý, công việc này thể hiện sự thừa nhận chuyển động. Chúng tôi hy vọng sẽ bao gồm nhiều asana yoga phức tạp hơn được hỗ trợ bởi AI trong tương lai.



6. TÀI LIỆU THAM KHẢO

- [1] Phân loại KNN [trực tuyến], Có sẵn: www.saedsayad.com/k_nearest_neighbors.htm
- [2] Cái túi - OpenPose [trực tuyến], Có sẵn: <https://www.geeksforgeeks.org/poseNetpose-Estimate/>
- [3] Mô hình phân loại KNN [trực tuyến], <https://www.pyimagesearch.com/2016/08/08/k-nn-classifier-for-image-classification/> Có sẵn
- [4] S. Patil, A. Pawar, A. Peshave, "Gia sư yoga: trực quan hóa và phân tích bằng thuật toán SURF". Trong: Kỷ yếu hội thảo nghiên cứu sau đại học về hệ thống điều khiển IEEE năm 2011, ICSGRC 2011.
- [5] Santosh Kumar Yadav, Amitojdeep Singh, Abhishek Gupta, Jagdish Lal Raheja, "Nhận dạng Yoga trong thời gian thực bằng cách sử dụng học sâu" (2019).
- [6] W. Wu, W. Yin, F. Guo, "Hệ thống chuyên gia tự học và tự hướng dẫn cho yoga", Proc. quốc tế Ứng dụng hệ thống làm việc thông minh, 2010.
- [7] H. -T. Chen, Yu-Zhen He, C. -L. Chou, S. -Y. Lee, B. -SP Lin và J. -Y. Yu, "Hệ thống tự đào tạo có sự hỗ trợ của máy tính dành cho bài tập thể thao sử dụng động cơ," 2013 IEEE International Conference on Multimedia and Expo Workshops (ICMEW), 2013, trang 1-4, doi: 10.1109/ICMEW.2013.6618307.
- [8] Shruti Kothari 2020 Phân loại tư thế Yoga bằng cách sử dụng Deep Learning Ph.D. luận án ĐẠI HỌC BANG SAN JOSE.
- [9] A. Chaudhari, O. Dalvi, O. Ramade và D. Ambawade, "Yog-Guru: Hệ thống chỉnh sửa tư thế yoga theo thời gian thực bằng các phương pháp học sâu," Hội nghị quốc tế về công nghệ máy tính và thông tin truyền thông năm 2021 (ICCICT), 2021, trang 1-6, doi: 10.1109/ICCICT50803.2021.9509937.
- [10] <https://janmesh18.github.io/Yoga.github.io/> -Truy cập trang web này để tham khảo và thực hiện Dự án