International Journal for Research in Engineering Application & Management (IJREAM)
ISSN : 2454-9150 Vol-10, Issue-01, Apr 2024

Yoga Pose Estimation Using YOLO Model

Mr. Yash Wanve, Student dept. of Information Technology SKN Sinhgad Institute of Technology & Science, Lonavala, Maharashtra E-mail: yashrwanve@gmail.com

Mr. Rohit Londhe, Student dept. of Information Technology SKN Sinhgad Institute of Technology & Science, Lonavala, Maharashtra E-mail: rohitlondhe2112@gmail.com

Mr. Uday Harle, Student dept. of Information Technology SKN Sinhgad Institute of Technology & Science, Lonavala, Maharashtra E-mail: udayharle233@gmail.com

Mr. Atharva Lohe, Student dept. of Information Technology SKN Sinhgad Institute of Technology & Science, Lonavala, Maharashtra, E-mail: atharva.lohe.al@gmail.com

Prof. P. S. Patil, Asst. Professor, dept. of Information Tech, SKN Sinhgad Institute of Technology & Science, Lonavala, Maharashtra, E-mail: pspatil.sknsits@sinhgad.edu

Abstract - In recent years, there has been a growing interest in using computer vision techniques in various fields, including fitness and wellness. This paper introduces a new method for estimating yoga poses using the YOLO (You Only Look Once) model. By taking advantage of YOLO's ability to process images quickly, we aim to accurately estimate yoga poses in real time during practice sessions. Our approach involves training the YOLO model on a custom dataset that includes annotated yoga poses. Through thorough testing, we show that our method is effective in identifying and categorizing yoga poses under different lighting conditions and body positions. Additionally, we explore integrating features that can provide immediate feedback to practitioners to help them improve their posture during yoga sessions. Our system is also capable of handling obstacles and changes in lighting, making it suitable for use in various settings such as yoga studios or homes. Overall, our research contributes to the advancement of computer vision in promoting health and wellness, providing a valuable tool for both yoga enthusiasts and instructors.

Keywords:* *Pose Estimation, YOLO, Key Points Detection, Virtual yoga, ML models

I. INTRODUCTION

Yoga, originating in ancient India, has become a worldwide phenomenon celebrated for its all-encompassing health advantages. The fusion of modern technology with traditional practices has reshaped how yoga is experienced. Nonetheless, accurately identifying its various poses presents a distinctive challenge for computer vision and artificial intelligence. Deep learning has introduced a new era of automated yoga pose recognition, streamlining the laborious manual assessment process.

Our research delves into this fascinating fusion, with a focus on utilizing the YOLO (You Only Look Once) pose model—an esteemed model renowned for its efficiency and precision—in the realm of yoga pose detection. This model's real-time capabilities offer automation and accuracy to yoga studios and practitioners globally, enabling instant feedback on posture alignment from both images and videos.

This study serves as a bridge between the physical practice of yoga and the digital world, addressing the specific requirements of the yoga community while highlighting AI's broader potential in enriching traditional practices. Through this exploration, we shed light on how modern technology can uphold and enhance ancient traditions.

Embark on this journey with us as we uncover the possibilities of Yoga Pose Estimation Using the YOLO pose Model, seamlessly amalgamating the past and the future.

II. Overview of YOLOv7 Model

YOLOv7, stands for You Only Look Once model version 7 is modern object detection algorithms. Developed as part of a tevolutionary series of YOLO fashions, YOLOv7 represents a massive bounce in actual-time computer vision tasks, promising unparalleled accuracy and performance in object detection. Unlike conventional Pose Estimation algorithms,Yolov7 pose is a high-level multi-person keypoint detector. It is much like the bottom-up method but

 DOI : 10.35291/2454-9150.2024.0184

heatmap free. It is an expansion of the single-frame pose detection model called YOLO-Pose. It combines the strengths of both top-down and bottom-up approaches. YOLOv7 Pose has been specifically trained using the COCO dataset, encompassing 17 key landmark configurations. It is applied in PyTorch making the code high-quality easy to personalize as in keeping with your want. The pre-trained keypoint detection model is **yolov7w6pose.Pth.**

III. LITERATURE SURVEY

A. Implementation of ML Technique for Identification of Yoga Poses:

The research paper titled "Implementation of Machine Learning Technique for Identification of Yoga Poses" authored by Yash Agrawal, Yash Shah, and Abhishek Sharma likely explores the use of machine learning techniques to automatically recognize and classify different yoga poses. In this paper, the authors likely discuss how machine learning algorithms are applied to analyze and identify various yoga postures.

The study involve the collection of image or sensor data from practitioners in different yoga poses and then using machine learning models to classify and recognize these poses accurately. This research could have implications for yoga practice, fitness tracking, and healthcare applications. [7]

B. Human Pose Estimation Using Convolutional Neural Network :

The research paper titled "Human Pose Estimation Using Convolutional Neural Network" probably focuses on using advanced computer algorithms called Convolutional Neural Networks (CNNs) to figure out how people are positioned in images or videos. This means it helps identify key points like the head, shoulders, elbows, hips, and knees in a person's body when they're doing different activities.

In the paper, it's likely that the authors discuss how CNNs, which are really good at analyzing images, are used to solve the challenges of figuring out how humans are moving and positioned. They might explain the techniques they used, how they collected data, the computer models they created, and the process they followed to get accurate and reliable results. [5]

This kind of research can be super useful in areas like computer vision , robotics , gesture recognition and human-computer interaction . Essentially, it helps computers understand how people move and stand, which can be used in many practical ways...

C. Infinity Yoga Tutor: Yoga Posture Detection And Correction:

The research paper titled "Infinity Yoga Tutor: Yoga Posture Detection and Correction" likely explores a approach to using technology for enhancing yoga practice. This paper may discuss the development and implementation of an innovative system or tool called "Infinity Yoga Tutor" designed to detect and correct yoga postures.

In all probability, the paper covers how this technology employs computer vision, machine learning, or sensor-based techniques to analyze the alignment and correctness of yoga poses performed by individuals. The system may provide real-time feedback to practitioners, helping them improve their form and reduce the risk of injury during yoga sessions.

This research has potential applications in the field of fitness and wellness, offering a digital solution to aid in yoga practice by making it more accessible and informative. The paper may delve into the technical aspects of this technology, its effectiveness, and its contribution to yoga education and training. [6]

IV. PREDICTION

In yoga pose estimation models, predictions are made based on input images or videos containing individuals performing yoga poses. The goal is to accurately identify and localize key points on the body corresponding to different parts of the yoga poses. Here's an explanation of how prediction works in a typical yoga pose estimation model:

Preprocessing**:** The input images or frames from videos are preprocessed to enhance features and remove noise. This may involve resizing, normalization, and possibly augmentation to improve model performance.

Feature Extraction**:** The preprocessed images are fed into the model, which typically consists of convolutional neural network (CNN) layers. These layers extract hierarchical features from the input images, capturing spatial information relevant to identifying body keypoints.

Keypoint Detection: The extracted features are then used to predict the coordinates of key points on the body associated with specific yoga poses. These keypoints may include locations such as wrists, elbows, shoulders, hips, knees, and ankles, depending on the pose.

Pose Estimation: Once the keypoints are detected, they are connected to form a skeletal representation of the body in the given yoga pose. This skeletal representation provides a visual understanding of the pose and facilitates pose classification.

Post-processing: The predicted keypoints and skeletal representations may undergo post-processing steps to refine the results. This could involve techniques such as non-maximum suppression to eliminate redundant keypoints and smoothing to reduce jitteriness in the predicted poses.

Output: The final output of the model typically includes the coordinates of the detected keypoints and possibly a confidence score indicating the model's certainty in its predictions for each keypoint. Output can be shown in Figure(1).

 DOI : 10.35291/2454-9150.2024.0184

ISSN : 2454-9150 Vol-10, Issue-01, Apr 2024

Visualization: The predicted keypoints and skeletal representations can be overlaid shown in Figure(1) onto the videos to provide a visual representation of the estimated yoga poses.

Figure[1]. Pose Prediction

V. CORRECTION

To implement pose correction using the YOLOv7 model, we can follow these steps:

Collect Correct Yoga Pose Images and Generate Dataset of Joint Angles: - Collect a dataset of correct yoga pose images where the poses are accurately executed.

- Utilize the YOLOv7 model to detect keypoints or joints of the body in these images.

- Calculate the joint angles (e.g., angles between limbs or body parts) based on the detected keypoints.

- Store this dataset with images and corresponding joint angles for training or reference.

Find Thresholds Using Lower Bound and Upper Bound:

- Define lower and upper bounds for acceptable joint angles based on the correct yoga pose dataset.

- These bounds represent the acceptable range of angles for each joint in a pose.

- For example, if a joint angle falls outside this range, it indicates a deviation from the correct pose.

Create a Function for Correction Using the Threshold Values:

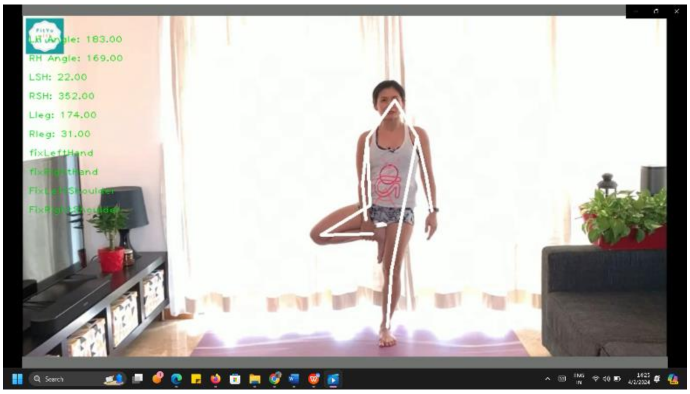
- Develop a function that takes the detected joint angles as input.

- Compare the detected angles with the predefined threshold values for each joint.

- If a detected angle exceeds the threshold range, apply correction to bring it within the acceptable range.

- The correction function can use various methods such as visual cues, auditory cues, or haptic feedback to guide the practitioner towards the correct pose.

As shown in Figure(2) which body part should be correct and which are already correct.



Figure[2]. Pose Correction

VI. MATH

In YOLOv7, the tanh function might be used within the neural network layers, particularly in the convolutional layers or fully connected layers, to introduce non-linearity and improve the model's ability to capture complex patterns in the input data.

tanh(x)=e^2x+1/e^2x−1

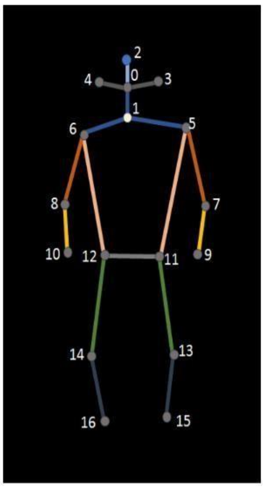
Mathematical Formula:

Detected Points={(x1,y1,c1),(x2,y2,c2),…,(xn,yn,cn)}

((xi,yi) represents the coordinates of the i-th key point on the image.

ci represents the confidence score of the i-th key point detection, usually ranging from 0 (no confidence) to 1 (full confidence).

The key points in human body can show in Figure[3].



Key Points	Skeleton
• Nose – 0	[15,13]
• Head bottom – 1	[13,11]
• Head top – 2	[16,14]
• Left ear – 3	[14,12]
• Right ear – 4	[11,12]
• Left Shoulder – 5	[5,11]
• Right shoulder – 6	[6,12]
• Left Elbow – 7	[5,6]
• Right Elbow – 8	[5,7]
• Left Wrist – 9	[6,8]
• Right Wrist – 10	[7,9]
• Left Hip – 11	[8,10]
• Right Hip – 12	[1,2]
• Left Knee – 13	[0,1]
• Right Knee – 14	[0,2]
• Left Ankle – 15	[1,3]
• Right Ankle – 16	[2,4]
	[3,5]
	[4,6]

Figure[3]- Key Point of Human Skeleton

VI. SUMMARY

In this study,we use the YOLOv7 pose pre-trained model to perform yoga pose estimation By utilizing the pretrained model, we were able to accurately extract keypoints of the human body, important markers on the human body, which

 DOI : 10.35291/2454-9150.2024.0184

are critical for yoga pose recognition. Our analysis focused on calculating angles between these keypoints, enabling us to recognize and categorize various yoga poses.

Results: The pre-trained model that we used has successfully identified key body points, allowing for precise angle calculation. The maccuracy is crucial in yoga, where the differences in body positioning can determine the correctness of a pose.

Limitations: as we are using pre-trained model for finding keypoints so accury of model is good but No model is perfect, and occasional mistake might occur, particularly in complex posesor with occluded body parts.

Future Research Directions

Real-Time Pose Estimation Optimization**:** Future research can be using this model on the real time webcam to enable realtime pose estimation. This could be crucial for applications where instant feedback is needed such as in live yoga classes or fitness apps.

VII. CONCLUSION

Our research focused on figuring out yoga poses using model called YOLOv7 Pose. We used pictures and videos as a input to find the angles and identify different yoga positions. This isn't just about technology it's about making yoga easier for everyone.By using ready-made tools, we made our work better and faster. We can now recognize yoga poses accurately from pictures and videos. This could help beginners and anyone practicing yoga to do it right.Our study shows how technology and yoga can work together. This model can make fitness easier for people. In the future we can add the features where model can recognize poses in real-time and we could make a app or website for virtual yoga. Our research isn't just about computers it is about making the world healthier and more connected. Our research has laid the groundwork for future innovations in health and fitness technology. By continuing to refine the model and explore new applications, we can help bridge the gap between traditional practices and modern advancements, creating a more connected, healthy world for all.

VIII. REFERENCES

[1] G. Chevalie, "LSTMs for Human Activity Recognition," 2016. [Online]. Available: https://github.com/guillaume-chevalier/LSTM-Human-Activity-Recognition

[2] H. Coskun, "Human Pose Estimation with CNNs and LSTMs," 2016.

[3] S.-H. Zhang et al., "Pose2Seg: Detection Free Human Instance Segmentation," 2018.

[4] Yoga CEERI, "Yoga Vid Collected", March 2019. [Online] Available: https://archive.org/details/Yoga VidCollected.

[5] B. Gil, "Single Pose Comparison," Becoming Human: Artificial Intelligence Magazine, 2017. [Online].

[6] S. E. Wei, V. Ramakrishna, T. Kanade, and Y. Sheikh, "Convolutionalpose machines," Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., vol. 2016-Decem, pp. 4724-4732, 2016.

[7] A. Hussain, E. Mkpojiogu and F. Kamal, Mobile Video Streaming.

[8] Ullah Khan and M. Ansari, "Performance Analysis of H.264 Video Coding Standard and H.263 Video Coding Standard", VSRD-TNTJ, vol. 2, no. 1, pp. 8-14, 2011.

[9] L. C Reddy and P. Hiremath, "RTSP Audio and Video Streaming for QoS in Wireless Mobile Devices", IJCSNS International Journal of Computer Science and Network Security, vol. 8, no. 1, 2008.

[10] S. Choudhary and P. Varshney, "A Study of Digital Video Compression Techniques", vol. 5, no. 4, 2016.

[11] U. Iqbal, A. Milan, J. Gall, "PoseTrack: Joint Multi-person Pose Estimation and Tracking," 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, 2017, pp. 4654-4663.

[12] P. Kulkarni, S. Mohan, S. Rogers, H. Tabkhi, "Key-Track: A Lightweight Scalable LSTM-based Pedestrian Tracker for Surveillance Systems," 2019, pp. 208-219.

[13] P. Kulkarni, S. Mohan, S. Rogers, H. Tabkhi, "Key-Track: A Lightweight Scalable LSTM-based Pedestrian Tracker for Surveillance Systems," 2019, pp. 208-219.

[14] A. Kendall, M. Grimes, R. Cipolla, "PoseNet: a convolutional network for real-time 6- DOF camera relocalization", IEEE Intl. Conf. Computer Vision, 2015.

[15] S. Kreiss, L. Bertoni, and A. Alahi, "PifPaf: composite fields for human pose estimation", IEEE Conf. Computer Vision and Pattern Recogn, 2019.

[16] P. Dar, "AI guardman – a machine learning application that uses pose estimation to detect shoplifters".

[17] D. Mehta, O. Sotnychenko, F. Mueller and W. Xu, "XNect: real-time multi-person 3D human pose estimation with a single RGB camera", ECCV, 2019.

[18] A. Lai, B. Reddy and B. Vlijmen, "Yog.ai: deep learning for yoga".

[19] M. Dantone, J. Gall, C. Leistner, "Human pose estimation using body parts dependent joint regressors", Proc. IEEE Conf. Computer Vision Pattern Recogn., 2013.

 DOI : 10.35291/2454-9150.2024.0184