AI-Assisted yoga Asanas in the future using Deep Learning and Posenet

Krupali Dhawale[1], Minakshi Ramteke[2] , Pranali Dhawas[3]
Madhuri Sahu[4]
[1,2,3,4]Assistant Professor ,Department of Artificial Intelligence ,G H Raisoni College of Engineering (GHRCE) Nagpur, India

---

Abstract - Yoga instructors have become teaching online due to the ongoing pandemic. Despite studying from top sources including videos, blogs, journals, and essays, users need live tracking to ensure proper posture and health. While technology can aid, beginner-level yoga practitioners rely on their teacher for identifying their proper posture.
The main objective of yoga pose detection and corrections is to deliver standard and precise poses for yoga using computer vision. If the yoga pose is not performed correctly, it may lead to severe injuries and long-term problems. Analysing human positions to identify and fix yoga positions can help humans live better lives in their own homes. Our work focuses on experimenting with different approaches to yoga position classification, therefore using PoseNet . Using such algorithms of deep learning, a person can determine the correct/ideal way/method to perform the particular yoga asana he or she is attempting to perform.

Key Words: Deep Learning, Posenet, Pose Estimation, posture recognition,tensor flow lite

1. INTRODUCTION

The rapid developments in information science and technology [1] have made living simpler than ever before. Everyone understands its importance in everyday life. The use of computers has had a significant impact on healthcare and other fields. Yoga, Zumba, martial arts, and other hobbies, along with normal medical procedures, are often regarded as effective ways to improve health. Yoga is an ancient Indian practice which encourages physical, mental, and spiritual health. Pose-Net artificial intelligence technologies [2], along with human posture assessment, may effectively integrate tech yoga. With much research, detecting human body postures remains challenging. To avoid injuries and stress through incorrect yoga postures, it's important to have a trainer. For those who cannot afford an instructor, AI can provide instruction. The instructor must adjust and perform each position for the individual before the practitioner can do the same. However, in today's world, more people have adapted to using the internet for what they need and feel comfortable doing so at home. Which requires system setup and camera. This is a basic yoga model that allows people to execute various yoga poses and adjust their yoga postures by looking at a reflected image[3][4].

In this paper we divided the model into three stages: 1. During the data collection/preparation stage, yoga postures were taken from open-source dataset containing seven positions ,YOGA -82 dataset and real images.

1. Poses are classified at multiple levels based on their visual appearance.
2. The feature extraction stage by using these collected data, the model undergoes

training in which the feature extraction done by PoseNet

2. RELATED WORK

Deep learning provides efficient evaluation of large data sets, making it a desirable research topic. Deep learning eliminates the need for feature extraction or design via recognizing complex trends in data and extracting features automatically. Earlier work focused on the outward appearance of structures, presenting human creation as a solid arrangement. Skeletons of humans are used for investigating the human body. The predefined mechanical body structure is commonly used to presume symmetry between body parts.[5] Recent advancements incorporate mixed, hierarchical, multimodal, and strong appearance models, combining falling and sequential prediction. The limbs are depicted.

Santosh Kumar Yadav et al. devised a method for identifying yoga poses using an RGB camera [5]. They used to collect the data, use a high-definition webcam. They used OpenPose to scan the user and identify crucial points. The time-distributed CNN layer recognized patterns between key points in a single frame, while the LSTM memorized patterns from recent frames. Their solution eliminates the need for Kinect or other specialized technologies to identify Yoga poses.

Ajay et al. [6] proposed a method that provides consistent feedback to practitioners, enabling them to detect correct and wrong stances. The researchers employed a data set of five yoga poses (Natarajasana, Trikonasana, Vrikshasana, Virbhadrasana, and Utkatasana) to train a deep learning model with a convolutional neural network for identifying yoga postures. This model identifies faults in posture and suggests corrections. It also classified the identified pose with 95% accuracy. This technology not only saves injuries but also enhances users' knowledge of yoga poses.

In their study "Yoga Pose Classification Using Deep Learning," Hang et al. [7] address a long-standing challenge in machine vision. Human activity analysis can aid various businesses, such as surveillance, forensics, assisted living, and home monitoring systems. In today's fast-paced world, many people prefer to exercise at home.

Patil et al. [8] developed a technique to recognize yoga posture differences between practitioners and specialists using image contours and SURF. Using contouring data simply for contrasting postures is insufficient.

Chaudhari et al. [9] developed a method that delivers explicit feedback to practitioners, allowing them to appropriately practice five yoga poses utilizing domain knowledge. A CNN model identified yoga poses, while a human-joint localization model detected pose flaws.

Guo et al. [10] designed an image- and text-based intelligent system for yoga training, integrating IMUs and tactors. However, this may be uncomfortable for users and disrupt the natural yoga stance. Additionally, the practitioner's posture was not considered.

Pose estimation from an image or video frame is a challenging problem. It is determined by the image's scale and resolution, as well as other factors such as lighting, variations, occultations, backdrop circumstances, and so on [11].

3. PROPOSED METHODOLOGY

Deep learning uses weights to mimic the workings of neurons in the human brain, making it a fundamental learning approach (ANN). This aids in visualizing neuronal connections. PoseNet:It is a computer vision model developed by Google that allows for real-time human stance assessment from photos or videos. It uses a convolutional neural network (CNN) architecture to recognize key points related to numerous body parts, including the nose, eyes, shoulders, elbows, wrists, hips, knees, and ankles. PoseNet is proficient in MobileNet Architecture. MobileNet is a Convolutional Neural Network developed by Google that was trained on the ImageNet dataset and is mostly used for image classification and target estimation. It is a lightweight model that use depthwise separable convolution to deepen the network while lowering parameters, computation costs, and increasing accuracy

PoseNet is trained to predict the places of these important elements, allowing it to assess an individual's pose inside an image or video frame. It can detect numerous persons in an image at once and returns a set of keypoints and a confidence score for each detected pose.
The model is lightweight and tuned for real-time performance, making it appropriate for a wide range of applications, including augmented reality, fitness tracking in fig 1.

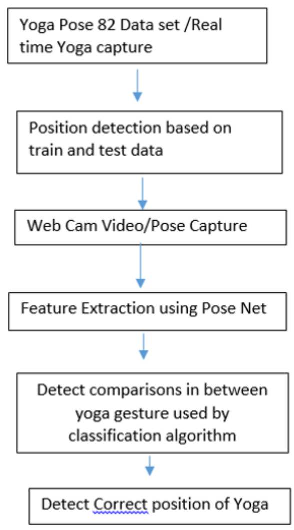

Fig1: Proposed Architecture for detecting YOGA Pose by using PoseNet Architecture

Layer of classification for Pose Net correction: The following steps are included in the joint estimating process:
1. Live video input: A live video feed of the user striking a yoga position serves as the system's input.
2. Preprocessing: The live video input is transformed into a picture, resized to a predetermined resolution, and its pixel values are normalized.

Joint Estimation: The updated TensorFlow Pose network is applied to the preprocessed live video in order to estimate the body joints' locations using kinematic representation. Data should be split into training and testing categories using Scikit-learn (Sklearn).
TensorFlow for Pose Correction Convolutional neural network (CNN) architecture is used by Movenet, a deep learning model for yoga posture detection. CNN creates a collection of 17 keypoint coordinates in the.csv file format that match the locations of body parts in an input image in the.jpg format. After that, the yoga posture being performed can be ascertained using these keypoint coordinates.

4. MATERIAL AND METHODS USING DATASET

The photos are from the Yoga-82 dataset, which was created exclusively for yoga pose classification tasks. It tries to collect more complicated and diversified data in order to create a more accurate and generalized system. 45 yoga pose sessions were chosen for this study from among the 82 accessible classes due to their difficulties.
The dataset used in this paper is one of the most recent and most challenging datasets, namely the Yoga82 hierarchically labelled dataset, which is used for yoga position estimation tasks with a wide range of complicated stance variants and photos acquired from real-world contexts. This paper introduces the work done

on a diverse variety of 11,000 photos split across 45 classes. The dataset was obtained via Kaggle [11] and other accessible sources. The illustrations depict five yoga poses: Utkata Konasana (goddess), Tadasana (mountain), Phalakasana (plank), Vrksasana (tree), and Virabhadrasana III (warrior 3). A total of 1578 images have been collected.

Label Hierarchy and Annotation:

Hierarchical annotations help users learn the network by providing detailed information about pose names, body postures (e.g., standing, sitting), spine effects (e.g., forward bend, back bend in wheel pose), and orientation (e.g., down-facing or up-facing).

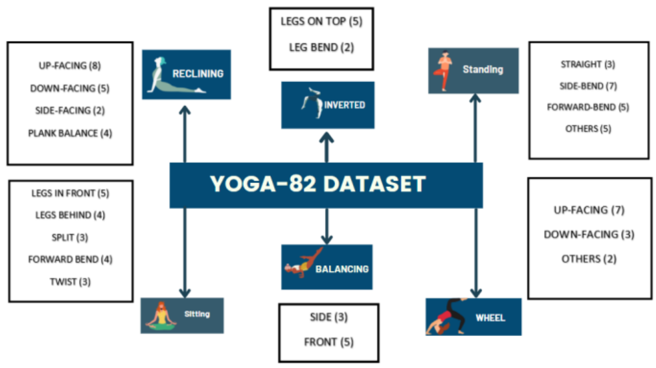

Fig-2: Label Hierarchy and Annotation

Feature Extraction:

The photos are manually labeled during pre-processing, followed by train-test validation and batch normalization to standardize the data. It has been significantly restructured and improved to meet the needs of the training model.

Pre-Trained Pose Net Model:

Pose estimate can be used to tackle some traditional problems, such as person counting in a frame, fall detection, smart fitness tracking apps, and so on. Essentially, posture estimation allows us to monitor human activity and make decisions. Previously, in the deep learning domain, feature extraction was done using HoG and SIFT-based approaches.

However, with CNN, these feature extraction processes have gotten more accurate when dealing with large amounts of data.

Fig- 3: Different Yoga Pose detection

PoseNet is used to find critical locations on human limbs. Features produce (x,y) coordinate values. Then, utilizing these keypoints, we may compute angles of various limbs of our body or use these points in a classifier model for human activity detection.

Proposed Model Performance

The model will make use of the user's webcam to extract the most important points from the yoga pose. This will classify it into one of the skilled yoga poses based on the feed, and we can watch an instructional clip of a yoga instructor correctly doing the posture that has been acknowledged, allowing them to learn and adjust the stance as necessary.

CONCLUSION

This research suggests developing a computer vision model for a yoga stance prediction job using a complex dataset with multiple pose variations. We proposed amodel by using Posenet to investigate human pose identification from a different perspective. By understanding body configurations in yoga postures, we can establish a labeling hierarchy.Video and image analysis can be used to verify the correct movement of yoga Asanas. Video-based analysis can benefit from models like Posenet, and KNN Classifier We describe a modified PoseNet architecture that uses hierarchy labels and improves performance compared classification.

REFERENCES

1. Shah, D., Rautela, V., & Sharma, C. (2021, September). Yoga pose detection using posenet and k-nn. In *2021 International conference on computing, communication and green engineering (CCGE)* (pp. 1-4). IEEE.

2. Upadhyay, A., Basha, N. K., & Ananthakrishnan, B. (2023, February). Deep learning-based yoga posture recognition using the Y_PN-MSSD model for yoga practitioners. In *Healthcare* (Vol. 11, No. 4, p. 609). MDPI.

3.Jadhav, R., Ligde, V., Malpani, R., Mane, P., & Borkar, S. (2023). Aasna: Kinematic Yoga Posture Detection And Correction System Using CNN. In *ITM Web of Conferences* (Vol. 56, p. 05007). EDP Sciences.

4. Verma, M., Kumawat, S., Nakashima, Y., & Raman, S. (2020). Yoga-82: a new dataset for fine-grained classification of human poses. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops* (pp. 1038-1039).

5. Yadav, S. K., Singh, A., Gupta, A., & Raheja, J. L. (2019). Real-time Yoga recognition using deep learning. *Neural computing and applications*, *31*, 9349-9361.

6. Chaudhari, A., Dalvi, O., Ramade, O., & Ambawade, D. (2021, June). Yog-guru: Real-time yoga pose correction system using deep learning methods. In *2021 International Conference on Communication information and Computing Technology (ICCICT)* (pp. 1-6). IEEE.

7.Zhang, Z., Tang, J., & Wu, G. (2019). Simple and lightweight human pose estimation. *arXiv preprint arXiv:1911.10346*.

8. Patil, S., Pawar, A., Peshave, A., Ansari, A. N., & Navada, A. (2011, June). Yoga tutor visualization and analysis using SURF algorithm. In *2011 IEEE control and system graduate research colloquium* (pp. 43-46). IEEE.

9. Chaudhari, A., Dalvi, O., Ramade, O., & Ambawade, D. (2021, June). Yog-guru: Real-time yoga pose correction system using deep learning methods. In *2021 International Conference on Communication information and Computing Technology (ICCICT)* (pp. 1-6). IEEE.

10.Wu, W., Yin, W., & Guo, F. (2010, May). Learning and self-instruction expert system for Yoga. In *2010 2nd international workshop on intelligent systems and applications* (pp. 1-4). IEEE.

11. Badiola-Bengoa, A., & Mendez-Zorrilla, A. (2021). A systematic review of the application of camera-based human pose estimation in the field of sport and physical exercise. *Sensors*, *21*(18), 5996.