IJARSCT

International Journal of Advanced Research in Science, Communication and Technology

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

ISSN: 2581-9429

Volume 5, Issue 6, May 2025

Impact Factor: 7.67

Yoga Pose Assessment Method Using Pose Detection for Deep Learning

Prof. J. R. Mahajan[1], Snehal Shelke[2], Sakshi Thombare[3], Sakhare Rohan[4], Kale Kaushal[5]

Guide, Department of Computer Engineering[1]

Students, Department of Computer Engineering[2,3,4,5]

Adsul's Technical Campus, Chas, India

Abstract: *An approach to accurately recognize various Yoga pose Assessment using deep learning algorithms has been presented in this work. In this system, we propose a Yoga pose assessment method using pose detection to help the self-learning of Yoga. The system first detects a Yoga pose using multi parts detection only with PC camera. In this system, we also propose an improved algorithm to calculate scores that can be applied to all poses. Our application is evaluated on different Yoga poses under different scenes, and its robustness is also. A hybrid deep learning model is proposed using convolutional neural network (CNN) and long short-term memory (LSTM) for Yoga recognition on real-time videos, where CNN layer is used to extract features from key-points of each frame obtained from Open-Pose and is followed by LSTM to give temporal predictions..*

Keywords: Yoga pose

I. INTRODUCTION

Human pose estimation is a challenging problem in the discipline of computer vision. It deals with localization of human joints in an images or video to form a skeletal representation. To variations, surroundings, and interaction of humans with the surroundings an application of pose estimation which has attracted many researchers in this field is exercise and fitness. One form of exercise with intricate postures is yoga which is an age-old exercise that started in India but is now famous worldwide because of its many spiritual, physical and mental benefits The problem with yoga however is that, just like any other exercise, it is of utmost importance to practice it correctly as any incorrect posture during a yoga session can be unproductive and possibly detrimental. This leads to the necessity of having an instructor to supervise the session and correct the individual's posture. Since not all users have access or resources to an instructor, an artificial intelligence-based application might be used to identify yoga poses and provide personalized feedback to help individuals improve their form. In recent years, human pose estimation has benefited greatly from deep learning and huge gains in performance have been achieved. Deep learning approaches provide a more straightforward way of mapping the structure instead of having to deal with the dependencies between structures manually. used deep learning to identify 5 exercise poses: pull up, Swiss ball hamstring curl, push up, cycling and walking. However, using this method for yoga poses is a relatively newer application. This project focuses on exploring the different approaches for yoga pose classification and seeks to attain insight into the following: what is pose estimation? What is deep learning? How can deep learning be applied to yoga pose classification in rea-time? This project uses references from conference proceedings, published papers, technical reports and journals.

II. MOTIVATION OF THE PROJECT

• Human pose estimation is a challenging problem in the discipline of computer vision.

• To automatically detect a person's pose in an image is a difficult task as it depends on a number of aspects such as scale and resolution of the image, illumination variation, background clutter, clothing variations, surroundings, and interaction of humans with the surroundings.

• There are a number of yoga asanas, and hence creating a pose estimation model that can be successful for all the asanas is a challenging problem.

• Yoga is widely recognized for its physical and and emotional benefits; however, the effectiveness of yoga largely depends on performing the poses correctly.

III. LITERATURE SURVEY

1. Implementation of Machine Learning Technique for Identification of Yoga Poses - Yash Agrawal, Yash Shah, Abhishek Sharma– In recent years, yoga has become part of life for many people across the world. Due to this there is the need of scientific analysis of y postures. It has been observed that pose detection techniques can be used to identify the postures and also to assist the people to perform yoga more accurately. Recognition of posture is a challenging task due to the lack availability of dataset and also to detect posture on real-time bases. To overcome this problem a large dataset has been created which contain at least 5500 images of ten different yoga pose and used a tf-pose estimation Algorithm which draws a skeleton of a human body on the real-time bases. Angles of the joints in the human body are extracted using the tf-pose skeleton and used them as a feature to implement various machine learning models. 80% of the dataset has been used for training purpose and 20% of the dataset has been used for testing. This dataset is tested on different Machine learning classification models and achieves an accuracy of 99.04% by using a Random Forest Classifier [1].

2. Yoga-82: A New Dataset for Fine-grained Classification of Human Poses - Manisha Verma, Sudhakar Kumawat, Yukta Nakashima – Human pose estimation is a well-known problem in computer vision to locate joint positions. Existing datasets for learning of poses are observed to be not challenging enough in terms of pose diversity, object occlusion and viewpoints. This makes the pose annotation process relatively simple and restricts the application of the models that have been trained on them. To handle more variety in human poses, we propose the concept of fine-grained hierarchical pose classification, in which we formulate the pose estimation as a classification task, and propose a dataset, Yoga-82§, for large- scale yoga pose recognition with 82 classes. Yoga82 consists of complex poses where fine annotations may not be possible. To resolve this, we provide hierarchical labels for yoga poses based on the body configuration of the pose. The dataset contains a three-level hierarchy including body positions, variations in body positions, and the actual pose names. We present the classification accuracy of the state-of-the-art convolutional neural network architectures on Yoga 82. We also present several hierarchical variants of Dense Net in order to utilize the hierarchical labels [2].

3. Recognition Of Yoga Poses Using Emg Signals from Lower Limb Muscles- Pradchaya Anantamek-: Exercise with yoga postures is very popular nowadays because yoga exercises can help to increase flexibility and muscle strength and improve the respiratory system. However, the correctness of the yoga postures is difficult to check, and thus practitioners may not be able to benefit from the exercises fully. This pa per presents a yoga posture recognition system to verify the correctness of the lower muscle movements while practicing yoga. The study included ten subjects, five males and five females. Data were collected during five yoga postures. This paper focuses on the use of Electromyography signals for analysing the motion of four lower-limb muscles of both legs. Recognition was performed with three machine learning algorithms. The results showed that the Random Forest Decision Tree algorithm has the highest accuracy in recognizing yoga postures in comparison with other algorithms and that the yoga posture recognition model is accurate at 87.43 percent [8].

4. Synthesizing Images of Humans in Unseen Poses- Guha Balakrishnan, Amy Zhao- We address the computational problem of novel human pose synthesis. Given an image of a person and a desired pose, we produce a depiction of that per son in that pose, retaining the appearance of both the person and background. We present a modular generative neural network that synthesizes unseen poses using training pairs of images and poses taken from human action videos. Our network separates a scene into different body part and background layers, moves body parts to new locations and refines their appearances, and composites the new foreground with a hole-filled background. These subtasks, implemented with separate modules, are trained jointly using only a single target image as a supervised label. We use an adversarial discriminator to force our network to synthesize realistic details conditioned on pose. We demonstrate image synthesis results on three action classes: golf, yoga/workouts and tennis, and show that our method produces accurate results within action classes as well as across action classes. Given a sequence of desired poses, we also produce coherent videos of actions [4].

5. Novel IoT-based Privacy-Preserving Yoga Posture Recognition System Using Low- Resolution Infrared Sensors and Deep Learning

-Munkhjargal Gochoo, Tan-Hsu Tan-: In recent years, the number of yoga practitioners has been drastically increased and there are more men and older people practice yoga than ever before. Internet of Things (IoT)- based yoga training system is needed for those who want to practice yoga at home. Some studies have proposed RGB/Kinect camera-based or wearable device- based yoga posture recognition methods with a high accuracy; however, the former has a privacy issue and the latter is impractical in the long-term application. Thus, this paper proposes an IoT-based privacy-preserving yoga posture recognition system employing a deep convolutional neural network (DCNN) and a low-resolution infrared sensor based wireless sensor network (WSN). The WSN has three nodes (x, y, and z-axes) where each integrates 8×8 pixels' thermal sensor module and a Wi-Fi module for connecting the deep learning server. We invited 18 volunteers to perform 26 yoga postures for two sessions each lasted for 20 s. First, recorded sessions are saved as .csv files, then pre-processed and converted to grayscale posture images. Totally, 93 200 posture images are employed for the validation of the proposed DCNN models. The tenfold cross validation results revealed that F1-scores of the models trained with xyz (all 3-axes) and y (only y-axis) posture images were 0.9989 and 0.9854, respectively. An average latency for a single posture image classification on the server was 107 ms. Thus, we conclude that the proposed IoT-based yoga posture recognition system has a great potential in the privacy- preserving yoga training system [7].

IV. PROBLEM DEFINITION AND OBJECTIVE

Yoga is a practice that emphasizes body posture, breathing techniques, and meditation. However, beginners and even experienced practitioners often face challenges in ensuring they are performing yoga poses correctly. Incorrect alignment of the body during yoga can lead to discomfort or even injury. In many cases, individuals rely on instructors for feedback, but in the absence of an instructor or for self-practice, it can be difficult to assess whether a pose is being performed correctly.

Objective

- Pose detection: Develop a system that can detect key body points (e.g., head, shoulders elbows, hips, knees, ankles) from input images or video streams.
- Yoga Pose Assessment: Implement a system that assesses whether a detected yoga pose is correct or incorrect by comparing the detected pose to an ideal reference pose.
- Feedback and Suggestions: Provide feedback to the user about the correctness of their pose. For example, if the knees are not aligned with the toes, inform the user and suggest how to correct the pose.
- User-Friendly Interface: Develop a simple and intuitive interface for users to easily assess their poses, whether through images or video (real- time).
- Performance and Accuracy: Ensure that the system works efficiently in real-time with minimal latency, especially if working with live video feeds.

V. FUNCTIONAL REQUIREMENTS:

System Feature

- To have understanding of the problem statement.
- To know what are the hardware and software requirements of proposed system.
- To have understanding of proposed system.
- To do planning various activates with the help of planner.

VI. NON-FUNCTIONAL REQUIREMENTS:

Performance Requirements:

The performance of the functions and every module must be well. The overall performance of the software will enable the users to work decently. Performance of encryption of data should be fast. Performance of the providing virtual environment should be fast safety requirements. The application is designed in modules is where errors can be detected and steadily. This makes it easier to install and update new functionality if re quired.

Safety Requirements:

The application is designed in modules where errors can be detected. This make is easier to install and update new functionality if required.

Security Requirements:

- Data Encryption: Requirement All sensitive user data, including personal information and session data, must be encrypted both in transit and at rest.
- User Authentication and Access Control: Users must be authenticated before accessing certain features of the system (e.g., saving session data, tracking progress).
- Secure User Data Storage: Any data that is stored on the system, such as user account details, pose history, or feedback, should be stored securely.
- User Consent and Privacy Policy: The system should request explicit consent from users for data collection and processing, and clearly explain how their data will be used.

VII. SYSTEM ARCHITECTURE:

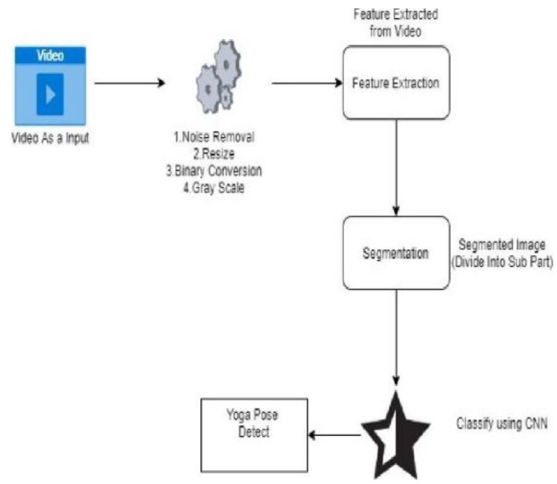


Fig.1: System Architecture

VIII. IMPLEMENTATION STEPS DATA COLLECTION:

1. Image capture: in the beginning stage, use an RGB camera to capture the image. The RGB camera is used to collect colour and depth pictures. The camera is installed and fixed on a tripod with a suitable frame that centres the person executing the yoga positions. The camera and the user are kept at a distance of 4 to 5 metres.

2. View Image: In the second step, use the function to take a sample image. OpenCV will be used to read the picture.

3. Carry out landmark detection: in the third step, a human skeleton of the human practicing the yoga positions is created using media pipe, and the findings are presented as 33 basic essential points. Following the pose identification, a total collection of thirty-three points identifying the main person's body joint positions in the image is generated. Each landmark contains: x: The picture width normalizes the landmark x-coordinate to [0.0, 1.0]. The picture height normalizes the landmark y- coordinate to [0.0, 1.0]. z: The z-coordinate of a landmark adjusted with the same level as x. It represents the depth of the landmark, with the origin being the halfway of the hips, therefore the lesser the number of z, the near the position to the camera.

I JARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 6, May 2025
Impact Factor: 7.67

IX. HUMAN POSE ESTIMATION:

Computer vision is used to estimate the human pose by identifying human joints as key points in images or videos, for example, the left shoulder, right knee, elbows, and wrist. Pose estimation tries to seek an exact pose in the space of all performed poses. It can be done by single pose or multi pose estimation: a single object is estimated by the single pose estimation method, and multiple objects are estimated by multi pose estimation. Human posture.

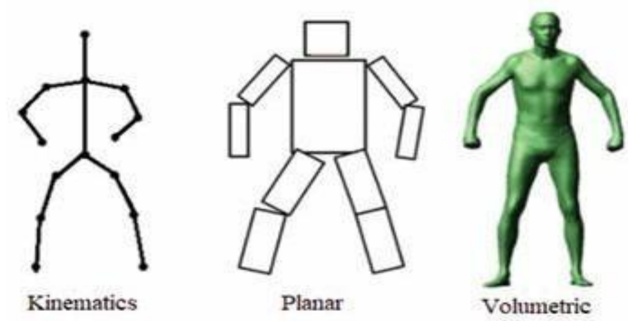


Fig.2: Human body modelling

assessment can be done by mathematical estimation called generative strategies, also pictorially named discriminative strategies. Image processing techniques use AI-based models, such as convolutional neural networks (CNNs) which can tailor the architecture suitable for human pose inference. An approach for pose estimation can be done either by bottom-up/top-down methods. In the bottom-up approach, body joints are first estimated and then grouped to form unique poses, whereas top-down methods first detect a boundary box and only then estimate body joints.

X. RESULTS

Pose estimation with deep learning

Deep learning solutions have shown better performance than classical computer vision methods in object detection. Therefore, deep learning techniques offer significant improvements in pose estimation. The pose estimation methods compared in this research include Epi polar Pose, Open Pose, Pose Net, and Media Pipe. Pose estimation for five yoga postures was done using different proposed techniques. The results of pose estimation were shown for each of the five asanas for all the four architectures used. For simplicity, the images of the same individual were shown (after taking consent) for all estimations and comparisons. The five yoga poses considered for posture estimation are as follows:

a. Ardha Chandrasana/half-moon pose,
b. Tadasana/mountain pose,
c. Trikonasana/triangular pose,
d. Veerabhadrasana/warrior pose
e. Vrukshasana/tree pose.

1. Epipolar Pose:

The Epipolar Pose constructs a 3D structure from a 2D image of a human pose. The main advantage of this architecture is that it does not require any ground truth data. A 2D image of the human pose is first captured, and then an epipolar geometry is utilized to train a 3D pose estimator. Its main disadvantage is requiring at least two cameras. The sequence of the steps for training is shown in Figure 2. The upper row of the (orange) depicts the inference pipeline and the bottom row (blue) shows the training pipeline. The input block consists of the images of the same scene (human pose) captured from two or more cameras. These images are then fed to a CNN pose estimator. The same set of images are then fed to the training pipeline, and after triangulation, the 3D human pose obtained (V) is fed back to the upper branch. Hence, this architecture is self-supervised.

Fig.3: Epipolar Pose. (a) Ardhachandr asana. (b) Tadasana. (c) Trikonasana. (d) Veerabhadrasana. (e) Vrukshasana

2. Open Pose:

The OpenPose is another 2D approach for pose estimation. Input images can also be sourced from a webcam or CCTV footage. The advantage of OpenPose is the simultaneous detection of body, facial, and limb key points. VGG-19, a trained CNN architecture from the Visual Geometry Group. It is used to classify images using deep learning. It has 16 convolutional layers along with 3 fully connected layers, altogether making 19 layers and the so-called VGG-19. The image extract of VGG- 19 is fed to a "two-branch multistage CNN,". The top part of predicts the position of the body parts, and the bottom part represents the prediction of affinity fields. i.e., the degree of association between different body parts. By these means, the human skeletons are evaluated in the image.

Fig.4: OpenPose. (a) Ardhachandrasana. (b) Tadasana. (c) Trikonasana. (d) Veerabhadrasana. (e) Vrukshasana

3. PoseNet:

The PoseNet can also take video inputs for pose estimation; it is invariant to image size; hence, it gives a correct estimation even if the image is expanded or contracted and can also estimate single or multiple poses. several layers with each layer having multiple units. The first layer includes input images to be analysed; the architecture consists of encoders that generate visual vectors from the image. These are then mapped onto a localization feature vector. Finally, two separated regression layers give the estimated pose.

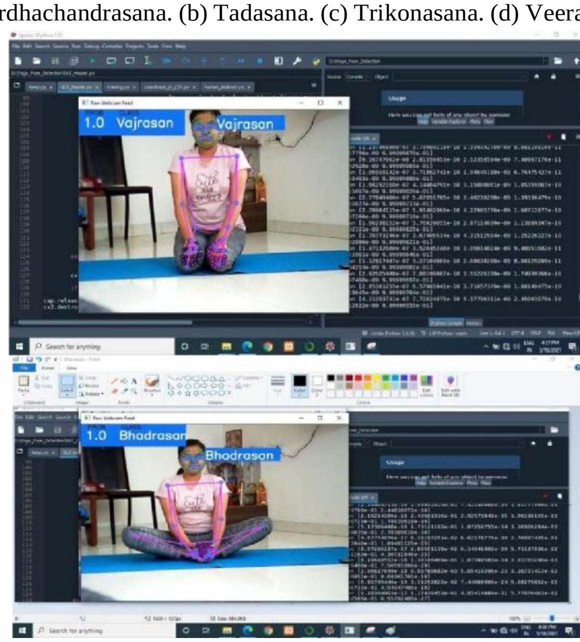
ISSN: 2581-9429
Volume 5, Issue 6, May 2025
Impact Factor: 7.67

Fig.5: PoseNet. (a) Ardhachandrasana. (b) Tadasana. (c) Trikonasana. (d) Veerabhadrasana. (e) Vrukshasana

4. MediaPipe:

This is an architecture for reliable pose estimation. It takes a color image and pinpoints 33 key points on the image. A two-step detector–tracker ML pipeline is used for pose estimation. Using a detector, this pipeline first locates the pose region- of-interest (ROI) within the frame. The tracker subsequently predicts all 33 pose key points from this ROI.

Fig.6: MediaPipe. (a) Ardhachandrasana. (b) Tadasana. (c) Trikonasana. (d) Veerabhadrasana. (e) Vrukshasana



DOI: 10.48175/568

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 6, May 2025
Impact Factor: 7.67

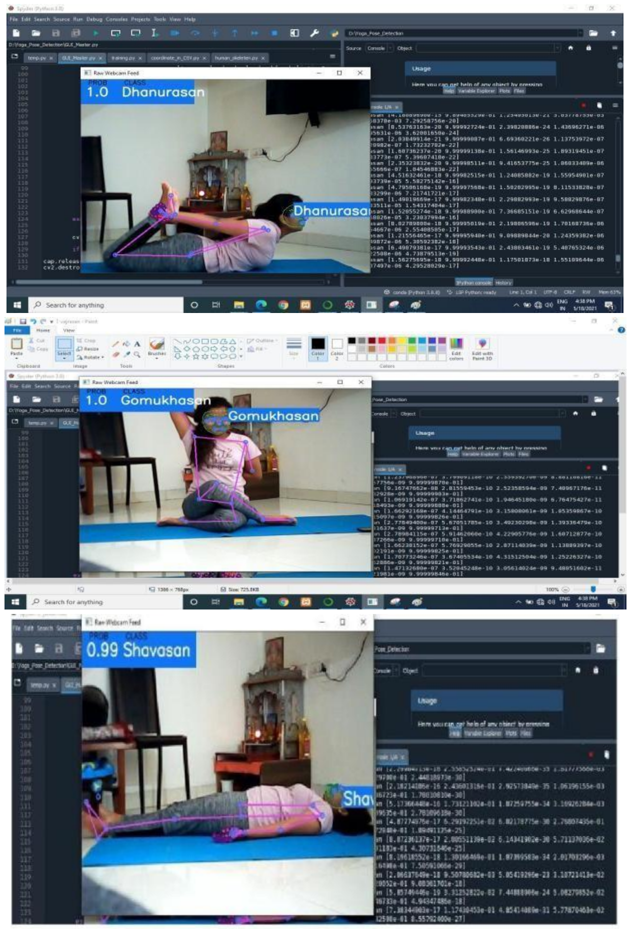


Fig.7. output

XI. CONCLUSION

We applied the time-distributed CNN layer to detect patterns between key points in a single frame and the LSTM to memorize the patterns found in the recent frames. Using LSTM for the memory of previous frames and polling for de-noising, the results make the system even more robust by minimizing the error due to false key point detection. Since the frames of a Yoga Images are sequential.

REFERENCES

[1]. Y. Agrawal, Y. Shah, A. Sharma– "Implementation of Machine Learning Technique for Identification of Yoga Poses," International conference on communication systems and network technology, 2020

[2]. M. Verma, S. Kumawat, Y. Nakashima "Yoga- 82: A New Dataset for Fine-grained Classification of Human Poses," Conference on computer vision and patterns recognition workshops, Cornell University, 2020

[3]. V. Akuhota and S. F. Nedler," Core Strengthening," American Academy of Physical Medicine and Rehabilitation, 2004.

[4]. Guha Balakrishnan, Amy Zhao G.Bradski and

a. Kaehler "Synthesizing Images of Humans in Unseen Poses",Learning OpenCV, O'Reilly, 2008.

[5]. Z. Cao, T. Simon, S.-E. Wei and Y. Sheikh," Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields," The Robotics Institute, Carnegie Mel lon University, 2017

ISSN: 2581-9429
Volume 5, Issue 6, May 2025
Impact Factor: 7.67

[6]. S.-E. Wei, V. Ramakhrisna, T. Kanade and Y. Sheikh," Convolutional Pose Machines," The Robotics Institute Carnegie Mellon University, 2016

[7]. Munkhjargal Gochoo, Tan-Hsu Tan "Novel IoT based Privacy-Preserving Yoga Posture Recognition System Using Low-Resolution Infrared Sensors and Deep Learning", 2020

[8]. Pradchaya Anantamek, "Recognition of Yoga Poses Using Emg Signals from Lower Limb Muscles", International ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunications Engineering (ECTI- NCON), 2019

[9]. P. K. Borkar, M. M. Pulinthitha and A. Pansare," Match Pose- A System for Comparing

DOI: 10.48175/568