

Improved Yoga Pose Detection Using MediaPipe and MoveNet in a Deep Learning Model

Deepak Parashar^{*ID}, Om Mishra^{ID}, Kanhaiya Sharma^{ID}, Amit Kukker^{ID}

Symbiosis Institute of Technology Pune, Symbiosis International (Deemed) University, Pune 412115, India

Corresponding Author Email: parashar.deepak08@gmail.com



<https://doi.org/10.18280/ria.370511>

ABSTRACT

Received: 13 May 2023

Revised: 20 July 2023

Accepted: 1 August 2023

Available online: 31 October 2023

Keywords:

yoga dataset, pose detection, image preprocessing, skeletonization, MediaPipe, MoveNet

The escalating global embrace of yoga as a holistic approach to well-being has accentuated the demand for refined and efficient techniques in yoga posture recognition. Traditional manual methods, although valuable, have exhibited protracted timelines and vulnerability to inaccuracies. In response, we introduce an innovative solution that harnesses the capabilities of deep learning (DL) models, elevating both the precision and accuracy of posture detection. Our approach predominantly leverages the Thunder variant of the MoveNet model, renowned for its exceptional proficiency in distinguishing an array of yoga postures. This model is seamlessly amalgamated with the MediaPipe technique, facilitating adept keypoint identification and skeletonization. In our proposed framework, input images undergo initial preprocessing, followed by skeletonization achieved through keypoint extraction. This pivotal process enables the encapsulation of distinctive points intrinsic to individual yoga poses. Central to our methodology is the incorporation of the large and diverse yoga (LDY) dataset, which encompasses five distinct yoga pose categories: Downdog, Goddess, Plank, Tree, and Warrior. A thorough evaluation demonstrates our approach's outstanding accuracy of 99.50% when deployed on the LDY dataset. As maintaining precise posture is pivotal in averting immediate discomfort and mitigating long-term health complexities, the implications of this advancement are profound. It charts a course toward more meticulous and accessible mechanisms for detecting yoga poses, thus profoundly influencing the physical and mental well-being of practitioners.

1. INTRODUCTION

In the contemporary environment, individuals are fixated on their routines due to the fast-paced nature of work. The rushed lifestyles that prevail have led to a multitude of severe health concerns among the majority. Yoga, an age-old communal endeavor rooted in India, encompasses aspects of mental, physical, and spiritual empowerment [1]. The significance of mental well-being spans across all stages of life, encompassing infancy, adolescence, and adulthood. Despite its undeniable importance, mental health often goes overlooked. The impact of mental health disorders reverberates on a global scale, affecting nearly 870 million individuals worldwide [2]. Since 2016, fitness Apps have seen a massive boom, with use increasing by more than 50% in only six months. The fitness App industry is expanding at a rate 85% higher than other App categories. Wearable technology's ability to assist users in keeping track of their fitness routines has contributed to the hike in popularity of yoga fitness applications [3].

Individual person behaviors are identified using human posture classification. Understanding human appearance and its related qualities is necessary for evaluating persons and uncovering human interactions with their environment, both of which are critical for industrial applications today [4]. Smartphones and other mobile devices have become more important in people's daily lives because of technology advancements that have made them more convenient. From an international perspective, it is predicted that by 2023, there will be 7.33 billion user increases from the anticipated 6.8

billion in 2018 [5]. Specifically, 90% of smartphone use time is spent interacting with mobile Apps. According to a recent systematic review, muscle strength, flexibility, balance, and quality of life are just some of the health-related important fitness metrics that have been shown to improve with yoga practice [6].

However, because of physiological variances, the review's findings cannot be extrapolated to all patients. Because of this, we decided to perform a systematic review to see how well yoga intervention fared with other exercise programs and inactive controls in improving health-related fitness and quality of life for individuals. The crucial information such as posture, attitude, gest [7]. Contend that the usage of ICTs and apps reduces the need for traditional modes of transportation as well as the need for physical venues to house genuine places of employment. Yoga, like any other exercise, must be done exactly since any wrong position during a yoga session can be useless and occasionally harmful [8].

The difference between an instructor's and a user has chosen body angles is subsequently calculated. If it is higher than the threshold, the procedure recommends that the part be repaired. Because of this concept, people will be able to practice yoga anywhere, including at home. As a result, everyone, regardless of age or health, may practice yoga. Hence, automated yoga pose detection is required for accurate and fast yoga practice [9].

Extensive research has been managed in the literature to construct several technology models. This section summarizes. Innovative techniques for human posture detection are linked.

Existing performance evaluation systems use simulators, sensors, and other sensing equipment and gadgets to detect how similar human body motions are to teacher actions. The system's ability to handle diverse self-taught scenarios and utilize Kinect for real-time recording of essential human body areas remains uncertain. As a result, it cannot recognize yoga poses. Despite the fact that the researchers described this approach for recognizing yoga poses, they did not teach the user how to fix improper yoga postures. As a result, it must be handled as well. In the study [10], they created an automated system-based yoga training approach for analyzing a practitioner's postures that consisted of feature points based on skeletons, contours, and dominating axes. The technology captures postural information and optimizes feature point identification, and supports axis building approaches. It creates yoga body skeletons by comparing performance with Microsoft Kinect and the open framework, as well as a camera, which is of RGB and a depth sensor to estimate/track the articulate stance [11].

Unfortunately, insufficient body map segmentation causes inconsistencies in offered instructions. They construct a computer using the capabilities of the Kinect gadget. Adaboost interface system for training purposes an algorithm was used to identify six yoga poses, it is effective in tracking depth, color, and body [12]. The intelligent system has also been built to recognize seated user in a wheelchair posture. After collecting data from a network of sensors using the neighborhood rule, the data is balanced using the k-nearest neighbor technique and the dimensions are reduced using principal component analysis. The pre-processed and balanced data is next subjected to the k-nearest neighbor approach [13].

The amount of data in this instance is a little smaller, but the outcome is astounding. To recognize posture, machine learning techniques such as convolution neural network (CNN), least square-support vector machine (LS-SVM) classifier and artificial neural network (ANN) are employed. The yoga data is acquired by inserting a sensing cushion consisting of a pressure sensor mat into the seat cushion of a children's chair and ten youngsters take five different viewpoints [14]. According to the constructability of inquiry signals from learned event dictionaries online, they provide a fully unsupervised dynamic coding strategy for recognizing strange occurrences in preprocessed images [14]. To categorize yoga positions, the algorithm employed deep learning techniques. Ancient machine learning methods need the extraction of characteristics as well as engineering, deep learning, algorithms on the other hand, reads data and extracts features. Using star skeleton computation method, a self-instructed system for the yoga posture is created for pose detection [15].

Drawing from the insights gained through the literature review, we can formulate the research problem in the following manner: In our fast-paced modern lives, we often become engrossed in daily tasks and overlook the importance of self-care, which encompasses physical and emotional well-being. Maintaining proper posture during exercise is essential, as improper posture can lead to physical issues. The consequences of physical inactivity extend beyond the physical realm, impacting mental health as well.

Inspired by cutting-edge methodologies, our proposed work aims to introduce a computer-aided automatic system for yoga pose detection. This innovative approach employs MediaPipe and a MoveNet-based deep learning technique, utilizing preprocessed input images capturing various yoga poses.

The main contribution of the proposed research work is as follows. (1) To propose an automated yoga pose detection framework using a novel MoveNet and MediaPipe-based deep learning model. (2) To enhance the performance of the model, keypoint-based skeletonization with the MediaPipe approach is used. (3) The Thunder version of MoveNet is used to improve the accuracy of yoga pose detection. (4) The effectiveness of the proposed method has been tested on a public yoga dataset of five classes.

The outlines of the remaining sections of the paper are as follows. The dataset used in the research work has been described in Section II. The proposed methodology is described in Section III. The experimental results and discussion is explained in Section IV. The conclusion is presented in Section V. In the end we present the references used in the article.

2. DATASET

The proposed method has been evaluated on large and diverse yoga (LDY) dataset consisting of five yoga classes. The yoga dataset used in this work is available on <https://www.kaggle.com/datasets/lakshmanarajak/yoga-dataset>. The dataset consists five yoga classes Downdog, Goddess, Plank, Tree, and Warrior. It consists of 2000 key pose images of these yoga classes [16]. These key pose images are separated for training and testing purposes. Different yoga pose images from the used LDY dataset is shown in Figure 1.

In this study, the large and diverse yoga (LDY) dataset has been used to evaluate the performance of the proposed model. A training dataset and a test dataset are separated in the available database. The image is stored in jpg file format having size of 300×300 pixels. Further, we use contrast histogram equalization strategy (CLAHE) to improve the pixel intensity and contrast of the images [17, 18]. The details of the training and testing images are given in Table 1.



Figure 1. Different yoga pose images from the used LDY dataset [16]

Table 1. Large and diverse yoga (LDY) dataset of five classes

| Yoga Class | Training Image | Test Image | Validation Image | Total |
|--------------|----------------|------------|------------------|-------------|
| Downdog | 300 | 100 | 28 | 428 |
| Goddess | 306 | 70 | 26 | 402 |
| Plank | 314 | 100 | 28 | 442 |
| Tree | 220 | 50 | 20 | 290 |
| Warrior | 321 | 90 | 27 | 438 |
| Total | 1461 | 410 | 129 | 2000 |

We create our training dataset by passing labeled images through all the models and then collecting all the landmark data and ground truth labels into CSV files. A dataset of yoga poses is used for further processing. A training dataset and a test dataset are separated in the directory. The dataset has been divided into train and test sets. We create a pose classification model that outputs the predicted labels and accepts the

landmark coordinates as input. With the use of landmark coordinates, our TensorFlow model can forecast the pose class that the subject in the input image would adopt. Two sub-models make up the main model, Sub model 1: computes a pose embedding (also known as a feature vector) using the coordinates of the detected landmarks as keypoints extracted from MediaPipe. Sub-model 2: predicts the pose class by feeding the pose embedding through several dense layers. Using the preprocessed dataset as a base, we trained the model. Fill the train and test datasets with the preprocessed CSV files.

3. PROPOSED METHODOLOGY

In the proposed framework, we introduced automatic detection of yoga poses using MediaPipe and a MoveNet-based deep learning approach. The research uses raw images of yoga poses as input data. The raw images are resized using the bicubic interpolation technique. Bicubic interpolation is a method to resize images while preserving their quality by calculating new pixel values based on the surrounding pixels. After resizing, the images are normalized. Normalization typically involves scaling pixel values to a specific range (e.g., 0 to 1) to ensure consistent data input to the model. In this case, the images are skeletonized using keypoints detected using the MediaPipe approach. MediaPipe is a framework developed by Google for building applications that involve perception tasks, like hand tracking, facial recognition, and pose estimation. Then, MediaPipe is employed to detect keypoints or landmarks on the yoga pose images. These keypoints are specific landmarks on the human body that help define the pose's structure and orientation. The skeletonized images, represented by keypoints, are fed into the MoveNet model. MoveNet is a deep learning model designed for human pose estimation. It's likely that the model can recognize and classify different yoga poses based on the key points. Then, the MoveNet model is utilized to classify the yoga poses. The keypoints provide a simplified representation of each pose, making it easier for the model to learn and differentiate between different poses.

3.1 Skeletonization using MediaPipe approach

In the proposed framework, the skeletonization is done through the keypoints detection using the MediaPipe approach. In this work, the traditional skeletonization technique is used by much more accurate and efficient deep learning approaches [10].

A deep learning and computer vision-based toolkit called MediaPipe, which is used to identify human skeletal posture. This demonstrates how keypoints are preferred to the real image for posture detection. To find human skeletal main points, MediaPipe architecture includes a detector and a tracker. While the tracker looks for posture landmarks, the detector locates the region of interest in the image [10, 15].

The framework by Google can detect 17 distinct points on the human body. There are alternatives to MediaPipe available like open pose, and blaze poses [7]. However, compared to others, MediaPipe, a cross-platform solution, is much faster in face, hand, and pose detection. MediaPipe demonstrates superior speed and compatibility with intricate perception pipelines compared to alternative methods [7, 10, 15] owing to its optimized algorithms and streamlined architecture. It is ideally suited for complex perception pipelines leveraging

accelerated inference [19]. Figure 2 show that posture detection MediaPipe skeletal view of different yoga poses after data pre-processing.

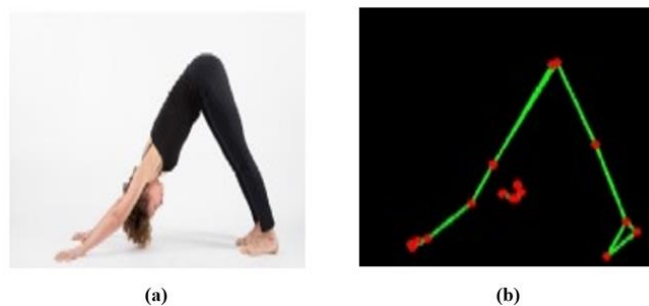


Figure 2. Posture detection MediaPipe skeletal view of different yoga poses after data pre-processing

3.2 MoveNet architecture

MoveNet employs a deep learning architecture designed for human pose estimation, which involves recognizing key points (also known as keypoints or landmarks) on the body to understand its pose. MoveNet uses a convolutional neural network (CNN) architecture that has been trained on a vast amount of annotated data containing human poses. This allows the model to learn the relationships between different body parts and their positions in various poses, this holistic approach enables MoveNet to accurately classify yoga poses based on their distinctive skeletal representations.

To analyze a person's quantitative accurately, the posture must be well described so that the numerical error may be estimated. The angle between certain bodily components is the most significant information for a yoga position [20]. Yoga asanas are easily identified in both live and recorded videos. The following processes are used by the system: Data is acquired from a camera or recorded footage in the first stage. Posture Estimation is a technique that captures a person's posture in coordinates, which are subsequently used to generate lines on detecting sensors. Coordinates may be used to create vectors, which can then be used to calculate the angles between two vectors. As a result, we describe someone's posture using many angles. This method makes use of a dataset to detect yoga positions. The fundamental function of this model is to find proper Yoga positions [21].

MoveNet is a quick and accurate model that recognizes 17 key points on the body. The model is offered in two varieties on TF Hub: Lightning and Thunder. Thunder is intended to be used for high-accuracy applications, whereas Lightning is intended to be used for low-latency applications. This is accomplished by running the model entirely within the browser, without the need to install any dependencies and with only one server call during the initial page load [22]. The MoveNet architecture is shown in Figure 3. The suggested system accurately distinguishes between the real and target locations and corrects the user. Because of technological advancements, everyone may choose the posture they wish to practice. To detect and fix stance difficulties, a real-time posture estimation approach is provided that leverages the TensorFlow model [23]. The model takes into account both the posture confidence score and the key points. The image is received by the model, which analyzes it to determine the posture. The method creates a collection of key points, and helps to find the confidence score and coordinates, there are

17 key points in each posture, each of which has an absolute x, y coordinate.

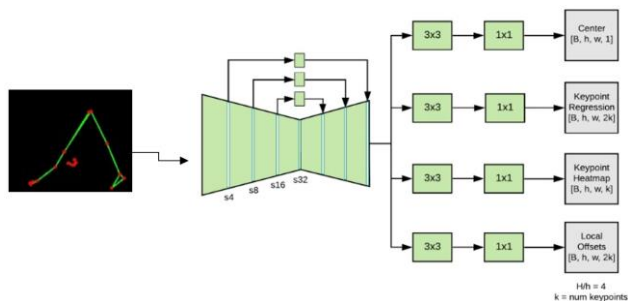


Figure 3. The proposed methodology based on MoveNet architecture

4. RESULT AND DISCUSSION

In this work, the input images are skeletonized and fed to the MoveNet model. In the proposed framework, the skeletonization is done through the keypoints detection using the MediaPipe approach. We used the Thunder version of MoveNet as it gives higher accuracy for a wide variety of yoga poses [24]. The study uses original yoga pose images as input. keypoints are identified using MediaPipe framework [25]. These keypoints guide the skeletonized images into the MoveNet model, designed for human pose recognition. The model likely identifies and classifies yoga poses based on these keypoints. Finally, MoveNet is employed to classify yoga poses, utilizing simplified pose representations from keypoints for detection and classification.

In this work, we used the large and diverse yoga (LDY) dataset consisting of five yoga classes. It consists of 2000 key pose images of these yoga classes. Where different classes of yoga Downdog, Goddess, Plank, Tree, and Warrior consists of 428, 402, 442, 290, and 438 images, respectively. These key pose images are separated for training and testing purposes. The original yoga images are resized into resolution of 256×256 pixels. Further, we apply CLAHE to improve the contrast and pixel intensity of the images. Then, the resized images are skeletonized using MediaPipe approach. In the proposed method, we choose 17 prominent keypoints on human body. The performance with different Keypoints is shown in Figure 4.

We tested the performance of the proposed method on different keypoints, where we obtained higher accuracy on 17 keypoints. The keypoints extracted from MediaPipe are localized using MoveNet model with the help of heatmaps. The person-centered heat map enhances model accuracy in yoga pose detection by precisely localizing keypoints for more precise pose estimation. The MoveNet model has two components feature extractor and prediction set heads. In this model, four prediction heads are used along with feature extractor for prediction. These heads are used to predict geometric center, full set of keypoints, and location of key points of skeletonized image. Although these predictions are computed concurrently, the following series of steps can be used to understand how the model works. (1) The centers of all the people in the frame are located using the person center heat-map, which is the arithmetic mean of all the key points that belong to a particular person. The region with the highest score is picked, with weighting based on distance from the

frame's center. (2) To create an initial set of key points for the person, slice the key point regression output from the pixel corresponding to the object center. Because this is a center-out prediction that must operate over multiple scales, the quality of the regressed key points won't be very accurate. (3) Key point pixels are scaled by weights inversely proportional to their regression distance from each key point. This ensures that we don't accept regressed key points from background participants because they normally won't be near the regressed key points and will therefore produce subpar outcomes. (4) To choose the final set of key point predictions, the locations with the highest heatmap values in each key point channel are obtained.

Figure 5 provides a plot of accuracy versus epoch. It can be observed from Figure 5 that the highest validation accuracy is obtained at epoch 27. Whereas the highest training accuracy is achieved at epoch 18. A plot of model loss function versus epoch for training and testing purpose is shown in Figure 6.



Figure 4. Plot of variation of performance with number of keypoints

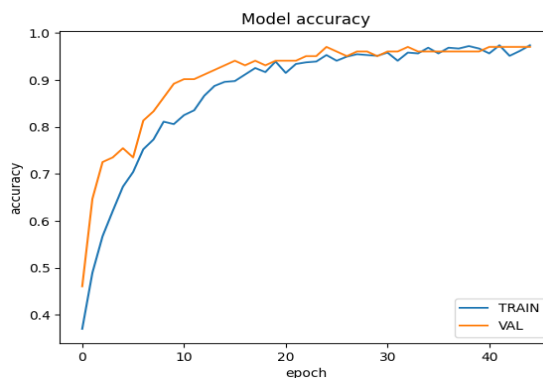


Figure 5. A plot of accuracy versus epoch

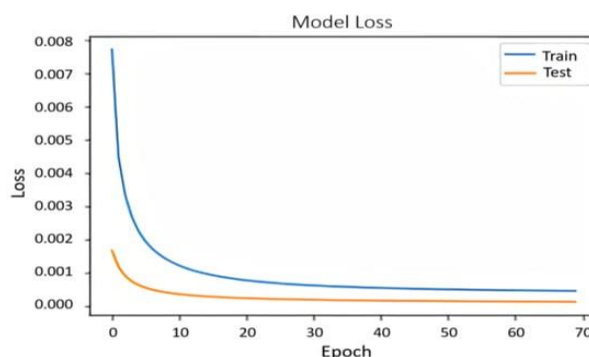


Figure 6. A plot of model loss function versus epoch

Table 2. Obtained confusion matrix for five classes of yoga

| | Downdog | Goddess | Plank | Tree | Warrior |
|---------|---------|---------|-------|------|---------|
| Downdog | 426 | 1 | 0 | 0 | 1 |
| Goddess | 1 | 399 | 2 | 0 | 0 |
| Plank | 0 | 1 | 440 | 1 | 0 |
| Tree | 0 | 0 | 0 | 289 | 1 |
| Warrior | 1 | 0 | 0 | 1 | 436 |

Table 3. The obtained ablation results of the proposed method on LDY dataset

| Dataset | Models/Approach | Ac (%) | Sn (%) | Sp (%) |
|------------------|---|--------|--------|--------|
| LDY yoga dataset | MoveNet (Non-MediaPipe and without Preprocessing) | 86.34 | 84.67 | 85.20 |
| | MoveNet (MediaPipe and without Preprocessing) | 89.50 | 87.67 | 88.10 |
| | MoveNet (Non-MediaPipe and with Preprocessing) | 92.34 | 89.67 | 87.20 |
| | MoveNet (MediaPipe and with Preprocessing) | 99.50 | 97.67 | 96.10 |

Table 4. Performance comparison with the existing methods for yoga pose detection

| Author | Year | Methodology | Ac (%) |
|------------------------|------|--------------------------------|--------------|
| Maddala et al. [9] | 2019 | Deep CNN and JADM | 87.27 |
| Shah et al. [13] | 2022 | PoseNet and KNN | 94 |
| Gupta and Jangid [11] | 2021 | Computer Vision | 97.4 |
| Proposed Method | 2023 | MoveNet and MediaPipe Approach | 99.50 |

The pose embedding is computed using the identified pose landmarks, and the pose class is predicted using our Keras model. In these, first we define the model to keep track of the checkpoint with the best validation accuracy, add a checkpoint callback. Then Start training to view your training history to determine overfit condition. Further, utilize the test dataset to assess the model. In which, the trained model, classify poses in the test dataset is used. To convert the prediction result to class name then evaluate the confusion matrix. The obtained confusion matrix for five classes of yoga is provided in Table 2.

In this study, parameter accuracy is used for performance measurement. The obtained results of the proposed method using different approach have been reported in Table 3. It can be observed from Table 3 that the proposed framework using MoveNet with the Non-MediaPipe approach without preprocessed yoga images obtained accuracy (Ac), sensitivity (Sn), and specificity (Sp) of 86.34%, 84.67%, and 85.20%, respectively. In another approach, we used MoveNet with MediaPipe approach without preprocessed yoga images, achieved accuracy, sensitivity, and specificity are 89.50%, 87.67%, and 88.10%, respectively. We used preprocessed yoga images using MoveNet with Non-MediaPipe, we got accuracy, sensitivity, and specificity are 92.34%, 89.67%, and 87.20%, respectively. Whereas, we obtained highest performance parameters using preprocessed images with MoveNet and MediaPipe approach, we achieved accuracy, sensitivity, and specificity are 99.50%, 97.67%, and 96.10%, respectively. We observed that the MediaPipe approach with preprocessed images can enhances the performance of the proposed framework.

The performance accuracy of the proposed method has been compared with the existing methods for yoga pose detection. The comparison with the existing method is shown in Table 4. The obtained accuracy of 87.27% using Deep CNN, Shah et al. [13] reported accuracy of 94% using PoseNet, and Gupta and Jangid [11] achieve 97.4% of accuracy using computer vision technique with traditional machine learning algorithms. Whereas we achieved far better accuracy of 99.50% using with MoveNet and MediaPipe approach. We choose four different models to test the robustness of the proposed method. The obtained results show that the proposed method using with MoveNet and MediaPipe approach outperformed to existing methods for yoga pose detection.

5. CONCLUSION

In this paper, we present a new automated method for yoga poses detection using MediaPipe and a MoveNet-based deep learning model. The performance of the proposed model involves employing the MediaPipe approach for key point-based skeletonization. Within the suggested framework, individuals' key points and skeletons are derived from the MediaPipe technique when processing input yoga images. Subsequently, the normalized and skeletonized images are supplied to the MoveNet model, enabling accurate detection of proper yoga poses. The LDY yoga dataset has been used to evaluate the performance of the proposed model. We achieved an accuracy of 99.50% using MoveNet's deep learning architecture. We chose four different models to test the robustness of the proposed method. The obtained results show that the proposed method using MoveNet and MediaPipe outperformed existing methods for yoga pose detection. In addition, we compared the results with the existing methods in terms of accuracy, and after comparison, we found that the proposed method outperformed for yoga pose detection. Yoga pose detection holds practical implications for personalized virtual yoga instruction, real-time pose correction, and progress tracking, fostering improved alignment and engagement in yoga practice.

CONFLICT OF INTEREST STATEMENT

All the Authors declare that there is no conflict of interest.

DATA AVAILABILITY STATEMENT

Our Manuscript has no associated data.

REFERENCES

- [1] Rajendran, A.K., Sethuraman, S.C. (2023). A survey on yogic posture recognition. *IEEE Access*, 11: 11183-11223. <https://doi.org/10.1109/ACCESS.2023.3240769>
- [2] Kim, Y.M., Son, Y., Kim, W., Jin, B., Yun, M.H. (2018). Classification of children's sitting postures using machine learning algorithms. *Applied Sciences*, 8(8): 1280. <https://doi.org/10.3390/app8081280>
- [3] 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. <https://doi.org/10.1007/s00521-019-04232-7>

- [4] Gochoo, M., Tan, T.H., Huang, S.C., Batjargal, T., Hsieh, J.W., Alnajjar, F.S., Chen, Y.F. (2019). Novel IoT-based privacy-preserving yoga posture recognition system using low-resolution infrared sensors and deep learning. *IEEE Internet of Things Journal*, 6(4): 7192-7200. <https://doi.org/10.1109/JIOT.2019.2915095>
- [5] Chen, H.T., He, Y.Z., Hsu, C.C. (2018). Computer-assisted yoga training system. *Multimedia Tools and Applications*, 77: 23969-23991. <https://doi.org/10.1007/s11042-018-5721-2>
- [6] Islam, M.U., Mahmud, H., Ashraf, F.B., Hossain, I., Hasan, M.K. (2017). Yoga posture recognition by detecting human joint points in real time using microsoft kinect. In *2017 IEEE Region 10 Humanitarian Technology Conference (R10-HTC)*, pp. 668-673. <https://doi.org/10.1109/R10-HTC.2017.8289047>
- [7] Ma, J., Ma, L., Ruan, W., Chen, H., Feng, J. (2022). A wushu posture recognition system based on MediaPipe. In *2022 2nd International Conference on Information Technology and Contemporary Sports (TCS)*. IEEE, pp. 10-13. <https://doi.org/10.1109/TCS56119.2022.9918744>
- [8] Chen, K.M., Tseng, W.S., Ting, L.F., Huang, G.F. (2007). Development and evaluation of a yoga exercise programme for older adults. *Journal of Advanced Nursing*, 57(4): 432-441. <https://doi.org/10.1111/j.1365-2648.2007.04115.x>
- [9] Maddala, T.K.K., Kishore, P.V.V., Eepuri, K.K., Dande, A.K. (2019). Yoganet: 3-d yoga asana recognition using joint angular displacement maps with convnets. *IEEE Transactions on Multimedia*, 21(10): 2492-2503. <https://doi.org/10.1109/TMM.2019.2904880>
- [10] Liu, J., Akhtar, N., Mian, A. (2020). Adversarial attack on skeleton-based human action recognition. *IEEE Transactions on Neural Networks and Learning Systems*, 33(4): 1609-1622. <https://doi.org/10.1109/TNNLS.2020.3043002>
- [11] Gupta, A., Jangid, A. (2021). Yoga pose detection and validation. In *2021 International Symposium of Asian Control Association on Intelligent Robotics and Industrial Automation (IRIA)*. IEEE, pp. 319-324. <https://doi.org/10.1109/IRIA53009.2021.9588714>
- [12] Trejo, E.W., Yuan, P. (2018). Recognition of yoga poses through an interactive system with Kinect based on confidence value. In *2018 3rd International Conference on Advanced Robotics and Mechatronics (ICARM)*. IEEE, pp. 606-611. <https://doi.org/10.1109/ICARM.2018.8610726>
- [13] Shah, D., Rautela, V., Sharma, C. (2021). Yoga pose detection using posenet and k-NN. In *2021 International Conference on Computing, Communication and Green Engineering (CCGE)*. IEEE, pp. 1-4. <https://doi.org/10.1109/CCGE50943.2021.9776451>
- [14] Mochiduki, S., Takahira, H., Yamada, M. (2015). Analysis of gazing points while viewing super-high-definition images at various viewing positions. In *2015 International Conference on Computer Application Technologies*. IEEE, pp. 54-57. <https://doi.org/10.1109/CCATS.2015.23>
- [15] Rajendran, A.K., Sethuraman, S.C. (2023). A survey on yogic posture recognition. *IEEE Access*, vol. 11, pp. 11183-11223. <https://doi.org/10.1109/ACCESS.2023.3240769>
- [16] 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. <https://doi.org/10.1109/CVPRW50498.2020.00527>
- [17] Parashar, D.R., Agarwal, D.K. (2021). SVM based supervised machine learning framework for glaucoma classification using retinal fundus images. In *2021 10th IEEE International Conference on Communication Systems and Network Technologies (CSNT)*, pp. 660-663. <https://doi.org/10.1109/CSNT51715.2021.9509708>
- [18] Parashar, D., Agrawal, D. (2020). Automated classification of glaucoma using retinal fundus images. In *2020 First IEEE International Conference on Measurement, Instrumentation, Control and Automation (ICMICA)*, pp. 1-6. <https://doi.org/10.1109/ICMICA48462.2020.9242702>
- [19] Wang, P., Wen, J., Si, C., Qian, Y., Wang, L. (2022). Contrast-reconstruction representation learning for self-supervised skeleton-based action recognition. *IEEE Transactions on Image Processing*, 31: 6224-6238. <https://doi.org/10.1109/TIP.2022.3207577>
- [20] Bajpai, R., Joshi, D. (2021). Movenet: A deep neural network for joint profile prediction across variable walking speeds and slopes. *IEEE Transactions on Instrumentation and Measurement*, 70: 1-11. <https://doi.org/10.1109/TIM.2021.3073720>
- [21] Cai, Z., Fernando, O.N.N., Ong, J.Y. (2022). PoseBuddy: Pose estimation workout mobile application. In *2022 International Conference on Cyberworlds (CW)*. IEEE, pp. 151-154. <https://doi.org/10.1109/CW55638.2022.00034>
- [22] Liu, D.X., Wu, X., Wang, C., Chen, C. (2017). Gait trajectory prediction for lower-limb exoskeleton based on Deep Spatial-Temporal Model (DSTM). In *2017 2nd International Conference on Advanced Robotics and Mechatronics (ICARM)*. IEEE, pp. 564-569. <https://doi.org/10.1109/ICARM.2017.8273224>
- [23] Suen, H.Y., Hung, K.E., Lin, C.L. (2019). TensorFlow-based automatic personality recognition used in asynchronous video interviews. *IEEE Access*, 7: 61018-61023. <https://doi.org/10.1109/ACCESS.2019.2902863>
- [24] Ünay, D., Stanciu, S.G. (2018). An evaluation on the robustness of five popular keypoint descriptors to image modifications specific to laser scanning microscopy. *IEEE Access*, 6: 40154-40164. <https://doi.org/10.1109/ACCESS.2018.2855264>
- [25] Gandhi, D., Shah, K., Chandane, M. (2022). Dynamic sign language recognition and emotion detection using mediapipe and deep learning. In *2022 13th International Conference on Computing Communication and Networking Technologies (ICCCNT)*. IEEE, pp. 1-7. <https://doi.org/10.1109/ICCCNT54827.2022.9984592>