International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 06 | June - 2023 SJIF Rating: 8.176 ISSN: 2582-3930

3D Yoga Pose Detection and Classification using Machine Learning Libraries

[1]Sarthak Mahendroo, [2]Shruti Gupta, [3]Harsh Rohilla, [4]Amit Yadav, [5]Ms. Gurpreet

[1,2,3,4] B.Tech. Students, CSE Department, HMRITM, New Delhi, India
[5] Assistant Professor, CSE Department, HMRITM, New Delhi, India

B. Tech 4th Year, Dept. of Computer Science and Engineering,
HMR Institute of Technology and Management, Hamidpur, Delhi-110036

ABSTRACT

India has a lengthy history of being connected to yoga, an age-old art form. It helps a person's body to be healthy and at the same time provides peace of mind**.** With the emergence of Covid-19, doing yoga in a class full of people has become difficult and can cause serious injuries if done without guidance on the other hand, here we identify the different yoga poses for a user to perform by developing a web application**.** The application uses open-source data containing images of 5 different yoga poses performed by different volunteers. OpenCV handles all the images for this application. The system has two phases, the first extracts points of data from the image dataset using the MediaPipe pose estimation library, the second phase pre-processes the acquired data with points and performs classification-based training and tests the data using machine learning algorithm. The machine learning algorithms used are Logistic Regression, Random Forest Classifier, Gradient Boosting Classifier, K-Nearest Neighbour Classifier, and Ridge Classifier**.** The system achieves an accuracy score of about 98%. The application is designed to process images with a threshold of still images and live video**,** so solutions below a certain score are unacceptable.

Human estimation is a difficult problem to solve in the domain of computer vision like locating human joints in an image or video and creating a skeletal representation. Automatic recognition of human pose in images relies on many aspects**, s**uch as image scaling and resolution, lighting changes, background noise, clothing changes, environment, and human- environment interactions, hence making it a difficult task**.** An application of human pose estimation that has intrigued many researchers in this field is exercise and fitness. This is an ancient practice that began in India, but is now world renowned for its many mental, physical and spiritual benefits. However, the problem with yoga, as with any exercise, is that improper posture during a Yoga session can be counterproductive and potentially harmful. And that's why you need an instructor to correct your posture. Not all users have access to instructors or instructing resources**.** However, one may use artificial intelligence-based application**s** to identify yoga poses and provide personalized feedback to improve personal modifications**.** In recent years, human pose estimation has benefited from deep learning, which has significantly improved performance. Machine learning approaches offer a simpler way to map structures than

dealing with dependencies between structures. Using machine learning and deep learning**,** we identified 5 yoga poses namely tree pose, plank pose, downward dog pose, warrior-2 pose and goddess pose.

Keywords: OpenCV, MediaPipe, Pose detection, Gradient boosting.

1. INTRODUCTION

Yoga is an art that originated in India a long time ago**.** It helps to improve a person's physical health and purifies a person's body, mind and soul**.** Yoga can cure many diseases without medicine in a world that invades everyone's mental peace, so yoga serves as the perfect recipe for this situation. All the systems need to evaluate a person's posture to detect yoga poses**.** People will benefit from this effort in doing this creative form of being fit more precisely and appropriately. The application aims to conduct a comparative examination of various yoga position classifiers. For the purpose of classifying yoga positions, we performed data collection and cleaning on a dataset of five poses. The model was created as a classification model for precisely predicting yoga poses. The proposed model has also undergone examination. The project's goal is to safeguard users against severe injuries that could result in chronic sickness. It contains simple self-guiding frameworks that make it possible for the user to do poses accurately and easily. Several machine learning methods have been used to classify and assess posture.

Human pose estimation is a computer vision task in which the pose of any person is detected when the camera takes an image of the person in front of it. Detection of human poses is done using key point detection method. These key points are the main points of the human body, including the nose, eyes, mouth, etc and various joints. Some of the key point identification techniques are considered are OpenPose, PoseNet, and PifPaf. These methods are employed to find different human poses. To deliver the required findings, they took into account variables like cross-entropy, LSTM (Long-Short Term Memory), perceptron, neural network, and convolution neural network. Moreover, there are two different ways to detect these key points**.**

Top-down approach: It is to find the bounding box that contains each person in the frame. Then**,** for each bounding box, the joint locations of the people in the box are determined. So, each bounding box has its own joints attached to it.

Bottom-up approach: This method is the antithesis of top-down strategy. In this first, all the joints that are available in the image are found, and then joints corresponding to every bounding box are separated to classify them according to each person.

The human posture estimation field is used in some significant work on yoga pose detection. We worked on a method that uses deep learning mainly convolutional neural networks for classifying yoga postures in images, the dataset employed consists of 1000 plus images distributed for 5 distinct yoga poses, with an accuracy rate of 85%. We then proposed a different method by finding the joint points of the human body and using them to create landmarks on the human body to estimate asanas or poses and accuracy is found out using various classification algorithms.

Mediapipe pose estimation: This is a pose estimation method developed by researchers of google and operates on the blaze fast model for the pose detection method. It is a fast model and performs at a 24FPS rate and hence is perfect for live video pose estimation.

Blazepose model returns 33 key points or landmarks from the given image in which a human is detected. These points are major joint points of a human body and the points returned are 3-D coordinates with a visibility value. The idea of Leonardo's Vitruvian Man is used to predict the coordinates of an invisible joint, from which the midpoint of a person's hip, the radius of a circle made up of humans, and the inclined line angle connecting the shoulder and hip's midpoint are all predicted. Classification in machine learning is the supervised machine learning methodology that works on obtaining certain values as output for several features. These output values are known as the target variables and their number is always greater than 1. The algorithms are mathematical functions that utilise the cost and error found in each piece of data we offer and train with a particular learning rate and gradient descent in order to decrease error. Examples of some classification-based machine learning algorithms include Logistic Regression, Random Forest Classifier, Gradient Boosting Classifier, K-Nearest Neighbour Classifier, and Ridge Classifier**.**

2. RELATED WORK

In the paper "A Computer vision-based Yoga Pose Grading Approach Using Contrastive Skeleton Feature Representations", for cross-validation of the training dataset, they used the cross-entropy loss approach and two benchmark datasets as well. Published in CVPR 2014, DeepPose was the first significant paper to apply deep learning towards the assessment of human posture. It achieved SOTA performance and outperformed earlier models back in 2014. In the paper "Real-time Yoga recognition using deep learning", they used CNN and LSTM for cross validation and deep learning as a data preparation method.

3. METHODOLOGY

The system comprises of two phases first is the training phase in which training of machine learning models happens and next is the testing phase in which testing of trained models and evaluation of their performances happen. Firstly, the environment for the system is created and all the necessary libraries are installed. The libraries that are necessary for the system are NumPy for numerical calculation, pandas for the dataset creation and data operations, seaborn for data visualization, sklearn for machine learning algorithms instances, OpenCV for computer vision tasks, MediaPipe for pose estimation.

After environment creation, data collection is done, hence open-source data is used from Kaggle where 5 yoga poses were performed by different volunteers both male and female. The data is in image format hence it is taken and fed to the next part of the system. Then the images are fed to the media pipe for pose estimation, it then detects the 33 key points on the given frames and provides the value of 3-D coordinates of these key points with the visibility value. These 33 key points form the basis of a new data set**,** and each key point coordinate and visibility become a function of the new data set. Then we apply data pre-processing methods to clean the data and make everything suitable for a machine learning model. Normalization is applied to the data set to pre-process the data**.** Normalization is the process of ensuring that all values in a data set are between 0 and 1**.** The main reason for this is that some machine learning models need normalized data to run

efficiently and give good results. Then, feature engineering is used to derive new features from the existing features in the dataset. To do this, we converted the key points into vectors**.** These vectors represent 3D body parts, and of these vectors are used to calculate joint angles. Finally, all processed data is passed to a classification-based machine learning algorithm for training. Test data must be acquired/generated first**.** This data must be new and should not consist of redundant data in the training phase.

We tried to compare various classifiers for the classification of yoga poses. These classifiers include Ridge Classifier, Logistic Regression Classifier, Gradient Boosting Classifier, Random Forest Classifier, and KNN Classifier. These methods of classification have aided in providing the best outcomes with increased accuracy. Web scraping has been used to collect the dataset. The aim was to collect the photographs in a way that would yield the best results. For each of the five poses that were assigned, the dataset was manually cleaned.

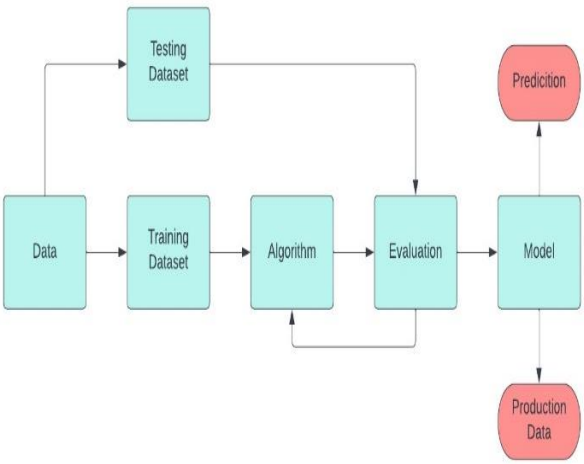

Fig. 1. The complete workflow of the project.

4. MODELING AND ANALYSIS

Model training: As the analysis of data is complete, modelling of data must be done using a classification-based machine learning algorithm. So, the system uses 5 different machine learning algorithms namely Logistic Regression, Random Forest classifier, Gradient Boosting classifier, K-Nearest Neighbour classifier, and Ridge classifier**.** Accuracy is a measurement criterion for machine learning model assessment.

A classification model has been created to accurately predict yoga positions. It has been suggested to evaluate the effectiveness of the categorization algorithms. For each approach we have used, it has been done with the aid of a confusion matrix. The project comprises a variety of aspects that have been broken down into various modules and components. The project was created as a result of the integration of these elements on multiple levels.

Collecting Dataset: We have acquired the dataset with the use of web scraping. To improve the dataset, it has been manually cleaned. If the image wasn't good enough, it's thrown out. The image is fed into the dataset if it meets the requirements. Dataset cleaning can also be done using a variety of methods.

Creating Landmarks: The execution of the algorithms has been marked by the creation of landmarks. With the help of MediaPipe, the predetermined landmarks have been occupied. The CSV file contains the landmarks if they were located. Using exception handling, it is possible to deal with missing landmarks.

Applying Algorithms: Logistic Regression, KNN Classifier, Gradient Boosting Classifier, Random Forest Classifier and Ridge Classifier are the five main sets of techniques that we have employed. To achieve the best

results, all these classification techniques were used when building the model. Each of these algorithms has been applied to the model separately. Further use has been made of the algorithm that displays the best results.

Best Model Selection: The best model has been chosen using a combination of many algorithmic parameters. It has produced a variety of results, and Gradient Boosting is the classification strategy that is most appropriate based on the best outcomes thus far.

Graphical User Interface: The GUI has been designed to be user-friendly. This makes it easier for the user to work on and explore the project. We have been able to give the users a faultless experience owing to the use of HTML, CSS, and Javascript.

There are three separate html files prepared since they will help in rendering the webcam feature and the contents simultaneously when the server starts.

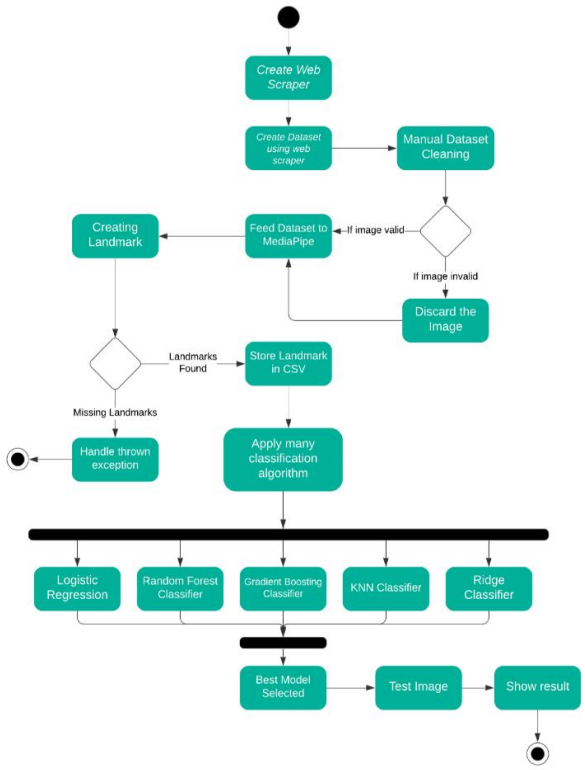

Fig. 2. Activity Diagram of the project.

5. RESULTS AND DISCUSSION

Accuracy measures the ability of the model to best predict successful classification in the total number of samples and accurately predict the target value. The precision and sensitivity we received were greater than what was formerly attained in the prior attempts by other people, which helped us solve the problem our application was seeking to tackle. We tested our data using a variety of algorithmic models to determine which was the best fit for the task and, Gradient Boosting Classification gave us an accuracy of almost 97%.

6. CONCLUSION

We started looking for a way to help people correct their posture while also creating a space where they could learn and grow in the convenience of their own homes, even if they didn't have access to or couldn't afford a personal trainer. And thanks to our efforts and good fortune, we were able to develop a platform where users can assess their posture as they go along and even improve it. Our model gives you an accurate reading of your posture for your reference, assisting you in correcting it and improving yourself with a simple user interface. In this study, a yoga pose classification application was successfully developed which works perfectly on images, static video, and live video of any user.

7. REFERENCES

1. Guddeti RR, Dang G, Williams MA, Alla VM (2018) Role of Yoga in cardiac disease and rehabilitation. J Cardiopulm Rehabil Prev. https://doi.org/10.1097/hcr.0000000000000372
2. Dr. Gaba, Importance of Yoga in the Pandemic, https://sarvodayahospital.com/blog/importance-ofyoga-during-this-pandemic
3. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3842656
4. Classification of yoga pose using machine learning techniques Author: J. Palanimeera, K. Ponmozhi https://doi.org/10.1016/j.matpr.2020.08.700
5. Ajeet Ram Pathak, Manjusha Pandey, Siddharth Rautaray, Application of Deep Learning for Object Detection, Procedia Computer Science, Volume 132, 2018, Pages 1706-1717, ISSN 1877-0509,https://doi.org/10.1016/j.procs.2018.05.144 (https://www.sciencedirect.com/science/article/pii/S1877050918308767).
6. Shruti Kothari 2020 Yoga Pose Classification Using Deep Learning Ph.D. thesis SAN JOSE STATE UNIVERSITY.
7. Y. Agrawal, Y. Shah, and A. Sharma, "Implementation of Machine Learning Technique for Identification of Yoga Pose," 2020 IEEE 9th International Conference on Communication Systems and Network Technologies (CSNT), 2020, pp. 40-43, DOI: 10.1109/CSNT48778.2020.911575
8. Narayanan, S. Sankara, and Misra, Devendra Kumar and Arora, Kartik and Rai, Harsh, Yoga Pose Detection Using Deep Learning Techniques (May 10, 2021). Proceedings of the International Conference on Innovative Computing & Communication (ICICC) 2021, Available at SSRN: https://ssrn.com/abstract=3842656 or http://dx.doi.org/10.2139/ssrn.3842656
9. Cramer, Jan Salomon. "The origins of logistic regression." (2002): 4 Bazarevsky, Valentin, et al. "BlazePose: On-device Real-time Body Pose tracking." arXiv preprint arXiv:2006.10204 (2020).
10. Kumar, Deepak & Sinha, Anurag. (2020). Yoga Pose Detection and Classification Using Deep Learning. International Journal of Scientific Research in Computer Science Engineering and Information Technology. 10.32628/CSEIT206623.
11. Nandakishore Joshi August 17, 2020, REALTIME 2D YOGA POSE ESTIMATION WITH CODE WALKTHROUGH, Medium, https://nandakishorej8.medium.com/realtime-2d-yoga-pose-estimationwith-code-walk-through-cfd69262d356
12. Syed Abdul Gaffar Shakhadri — May 18, 2021, Pose Estimation using OpenCV, Analytics Vidhya, https://www.analyticsvidhya.com/blog/2021/05/pose-estimation-using-opencv/
13. https://www.hindawi.com/journals/cin/2022/4311350/
14. C. C. Hsieh, B. S. Wu, and C. C. Lee, “A distance computer vision assisted yoga learning

system," Journal of Computers, vol. 6, no. 11, pp. 2382–2388, 2011.View at: Google Scholar.

15. Pauzi A.S.B. et al. (2021) Movement Estimation Using Mediapipe BlazePose. In: Badioze Zaman H. et al. (eds) Advances in Visual Informatics. IVIC 2021. Lecture Notes in Computer Science, vol 13051. Springer, Cham. https://doi.org/10.1007/978-3-030-90235-3_49
16. M. T. Uddin and M. A. Uddiny, "Human activity recognition from wearable sensors using extremely randomized trees," in Proceedings of the 2015 International Conference on Electrical Engineering and Information Communication Technology (ICEEICT), pp. 1–6, IEEE, London, UK, 2015 May.View at: Google Scholar.