YOGA POSE DETECTION USING MACHINE LEARNING LIBRARIES

Kushagra Anand[1], Kartik Verma[2], Shubhum Verma[3], Ms. Deepti Gupta[4] (Guide), Ms. Kavita Saxena[5] (Co-Guide)

[1]Kushagra Anand, Dept. of Computer Science Engineering, Maharaja Agrasen Institute of Technology, Delhi, India
[2]Kartik Vermaq, Dept. of Computer Science Engineering, Maharaja Agrasen Institute of Technology, Delhi, India
[3]Shubham Verma, Dept. of Computer Science Engineering, Maharaja Agrasen Institute of Technology, Delhi, India
[4]Ms.Deepti Gupta, Dept. of Computer Science Engineering, Maharaja Agrasen Institute of Technology, Delhi, India
[5]Ms.Kavita Saxena, , Dept. of Computer Science Engineering, Maharaja Agrasen Institute of Technology, Delhi,India

---

Abstract - Over many years yoga has been the source of energy and enlightment. The existence of yoga traces back to India, over 5000 years ago. Yoga is derived from a Sankrit root word "Yuj" which means "to join", "to yoke" or "to unite". All through the years yoga has been practiced for various reasons. Our aim is to come up with solutions of how to help people in the formation of yoga poses with accuracy. Along with this our focus is to solve the problem of having an instructor, in order to perform yoga. The project aims at designing a system that can detect a yoga performer's pose in real-time and predict whether he/she is doing it correctly or not. We plan on accurately predicting five popular yoga poses namely, downward dog, plank pose, tree pose, goddess pose and warrior-2 pose. This would involve training a Machine Learning model to complete the required task. OpenCV would be used for handling the images and pose detection will be performed using MediaPipe. A web application will be developed providing the user with a platform to easily perform his task with ease.

Key Words: OpenCV, MediaPipe, Pose detection, Gradient boosting

1. INTRODUCTION

Yoga is a 5000-year-old science covered in the views of spirituality, which involves physical mental and spiritual practices aiming at providing stability and calmness to the mind. But to reap maximum benefits, one needs to ensure that it is done properly. Although it is an old science but has gained popularity recently as a new way to find peace and bring harmony to both body and mind.Yoga can help in all-round development of the mind and body but the postures the processes play an important role because incorrect postures lead to serious and fatal injuries. In the world of technology 2 where the information is just a click away, people tend to follow yoga videos. With these videos, people usually tend to neglect the nuances of performing a posture. There are certain postures that needs analysis and accuracy. Therefore, it is necessary to adopt the right posture with accuracy. Our yoga pose detection model from video plays a critical role in various applications such as quantifying physical exercises, sign language recognition, and full-body gesture control. The purpose of this project is to design a system that predicts the yoga pose a person is performing and ensuring that he/she is doing it correctly. It is aimed towards people who're new to yoga and are aiming to perform the poses with high accuracy. This might be useful for trainers who can't individually pay attention to each new trainee, allowing them to carefully monitor performance of multiple trainees without paying close attention to each of them and this can also be used by individuals for monitoring their own posture and performance and help them in improving themselves. This project will help people out in doing this artistic form of being fit more accurately and correctly. The project aims at comparative analysis of different classifiers for Yoga pose classification. In this project we have done collection and cleaning of dataset of five poses for yoga pose classification. The model has been designed as a classification model for predicting yoga pose accurately. The evaluation of the proposed model has also been made. The aim of the project is to protect the user from major injuries that may lead to chronic illness. It has easy to use self-guiding frameworks that help the user to perform pose with ease and accuracy. Posture assessment & classification has been made using various machine learning algorithms.

2. Related Work

In "A Computer Vision-Based Yoga Pose Grading Approach Using Contrastive Skeleton Feature Representations" paper they used cross entropy loss method for cross validation of training dataset and they also used two benchmark dataset. DeepPose was the first major paper to apply Deep Learning to human posture assessment, published in CVPR 2014. Back in 2014, it attained SOTA performance and outperformed previous models. In "Real-time Yoga recognition using deep learning" they use deep learning as a data preprocessing technique and for cross validation they use CNN and LSTM.

3. Literature Survey

The model is trained to classify poses. Extensive research has been conducted in recent years to provide results that accurately classify the three poses. The studies in these studies relate to the recognition and classification of yoga poses. Important point detection methods used in these studies include OpenPose, PoseNet and PifPaf. These

methods are used to detect various human poses. We consider factors such as cross-entropy, LSTM (Long ShortTermMemory), perceptrons, neural network s, and neural convolutional networks to deliver the results you need. Previous limitations in research include that works containing key points are not scaled well and are limited to finding patterns in various human poses. It has many applications, such as software engineering. In recent years, they have initiated processes and models for implementing projects for sustainable human development and longevity. Some used random forests to detect human activity using sensors. Detect human activity using hidden Markov models and detected body parts. We used this method to detect activities in 6 houses and achieved an accuracy of 97.16%. This method is used to monitor servicesin smart homes. Ituses ambient background noise to detect human activity using wearable sensors that detect sound, achieving 96% accuracy.accuracy in real time for a set of 12 different people showing its ability to perform well with new sub-jects and conditions. Significant work has also been done in developing automated systems, which Analyze yoga and sports activities. The Speeded Up Robust Features (SURF) Algorithm is an automated system for yoga for novice users, and comparison with yoga videos from experts may not be enough. Detect yoga poses for an automated project using AdaBoost classifiers and motion sensors with 94.8% accuracy. Another system presents him with 3 yogaposes with an accuracy of 82.84%. The system used deep learning techniques to classify yoga poses. In traditional machine learning, models require extracted features and engineering, whereas deep learning understands data and extracts features. A self-taught system for yoga poses is created using star skeleton calculations. Kinect is used to extract body contours from the user's body map, achieving 99.33% accuracy. We used hash-based learning to extract human posture from pressure sensors. Since these sensors are not always portable, they are not used in the proposed system. A hybrid model CNN with LSTM is used to estimate the poses used in OpenPose to classify yoga poses, and this model includes feature extraction. Comparisons are made between deep learning and machine learning models, and between CNN and hybrid models.

4. DATASET

The dataset has been created in such a way that it is helpful for the training and testing of the models. The dataset has been collected with the help of web scrapping. Web 4 Scrapping scripts have been built to scrap images as per expectations. A web scraper was designed to create and obtain a collection of images that has formed the dataset. This has been separated into five folds as per the requirements of the project. The dataset so gathered need cleaning. The cleaning hasbeen done manually.

Fig-1: Normal Image

Fig-2: Yoga Pose Detected

4.METHODOLOGY

A dataset of 5 poses using web scraping was implemented. Once cleaned and processed, the data set is ready for further use. The data set is loaded and the images are converted to the desired format. Mediapipe was used to determine the number of landmarks on the human body. Based on these results, i.

H. The positions of landmarks so formed are used to determine the formed pose. Used for model training.

• The classification techniques used in the project are as follows.

• Logistic regression, statistical analysis (also known as logit model) is widely used for predictive analytics and modeling and extends to applications in ML. In this analytical approach, the dependent variable can be finite or categorical, A or B (binomial regression) or a set of finite options A, B, C, or D (multinomial regression). It is used in statistical software to understand the relationship between a dependent variable and one or more independent variables by estimating probabilitiesusing logisticregression equations.

• A random forest classifier consists of a large number of individual decision trees acting as an ensemble. Every tree in a random forest spits out a class

• A Gradient Boosting Classifier is a special kind of ensemble learning technique that works by combining multiple weak learners (low-accuracy

predictors) into one strong learner (high-accuracy model). It works by making each model pay attention to the errors of its predecessor.

• ANN Classifier is a supervised ML algorithm that assumes similarities exist in close proximity. Predicts the class or continuous value of a new data point given its K nearest neighbors (datapoints).

•Ridge Classifieris basedon the Ridge Regression metho d, transforming the label data to [-1, 1] and solving the problem with a regression method.

The highest predicted value is accepted as the target class and multi-output regression is applied to multiclass data.

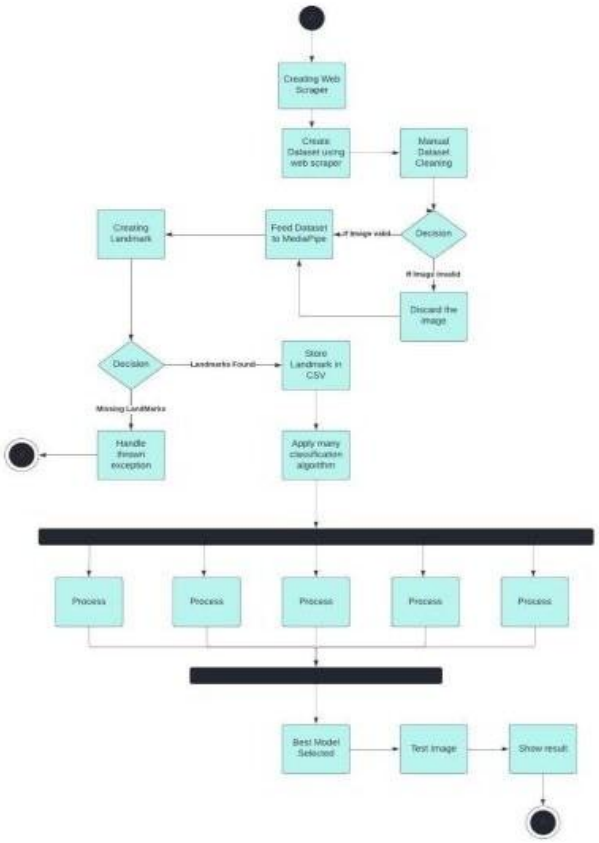


Fig. 3. It showcases the activity diagram of the project. It includes the actual workflow of the project. It has explained the project in anelaborate way with all the components and and modules.

5. Implementation

The main aim of the project is to design a system that ca detects yoga pose of the performers in real time. Along with real time prediction the project is able to predict whether the user is doing it correctly or not. Our aim is to predict five major yoga poses with accuracy along with providing accuracy with pose detection. The poses: Downward Dog, Plank Tree, Goddess Pose and Warrior-2 pose are the poses that we 6 are focusing at. Our project will help its users to learn yoga with ease andimprove his/her lifestyle, without the interaction or concern of the third person. Our project is extremely beneficial as it ensures proper social distancing, especially in the times of the pandemic. We have initiated comparative analysis for Yoga poses classification, of different classifiers. These classifiers are: Logistic Regression, Random Forest Classifier, Gradient Boosting, KNN Classifier, Ridge Classifier. Theseclassification techniques have helped give the best results with greater accuracy. The dataset has been gathering with the help of web scrapping. The task was to gather the images in such a way that it provides best outcomes. The cleaning of the dataset has been done manually for all the five poses, so classified.

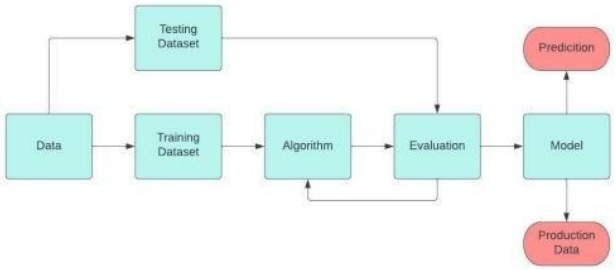


Fig. 4. It showcases the complete workflow of the project. It has the involvement of testing ang training dataset brief description.

Classification model has been designed for predicting yoga poses with precision. The evaluation of theperformance of the classification algorithms has been proposed. It has been done with the help of confusion matrix for every algorithm that we have implemented. There are different facets of the project which has beenclassified into different modules and components. The integration of these components at various levels have lead to the creation of the project.

• Collecting Dataset : As mention in the section(Dataset),we have gathered dataset with the help pf web scrapping. The dataset has been cleaned manually to improve the dataset. If the image was not up to the mark, the image is discarded. If the image 7 is fulfilling the criteria it is fed to the dataset. Various techniques can also be used for dataset cleaning.

• Creating Landmarks: Landmarks have been created for the implementation of the algorithms. The set of predefined landmarks has been occupied with the help of MediaPipe. If the landmarks were found they arestores in the CSV file. If there are some missing landmarks, it is handled with the help of exception handling.

• Applying Algorithms: 5 major sets of algorithms that we have used are : Logistic Regression, Random Forest Classifier, Gradient Boosting, KNN Classifier, Ridge Classifier. All these classification algorithms have been implemented while creating the model in order to get the best results. These algorithms have been implemented

individually over the model. The algorithm that showcases the best results have been used further.

• Best Model Selection: Best model has been selected based on the basis of different parameters of different algorithms. It has provided different results and based on the best results so far the classification technique best suited is Gradient Boosting.

• Graphic User Interface: GUI has been created in the user-friendly manner. This helps the user to work and explore the project at an ease. The use of HTML,CSS and Javascript has helped us provide a flawless experience to the users. HTML component contains 3 files: base.html, index.html and webcam.html. They help to render the contents when the server starts along with rendering the webcam function at the time of execution.

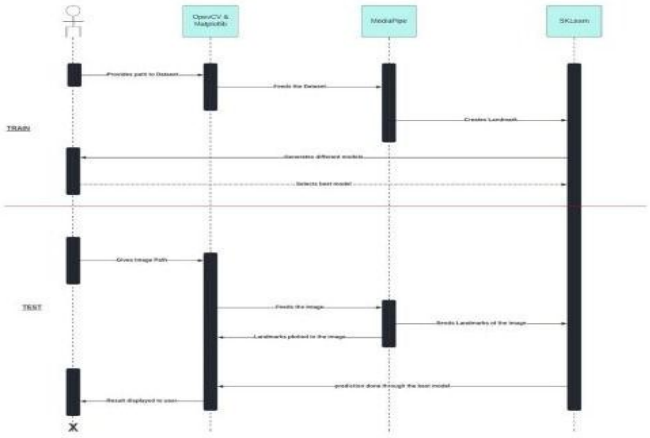


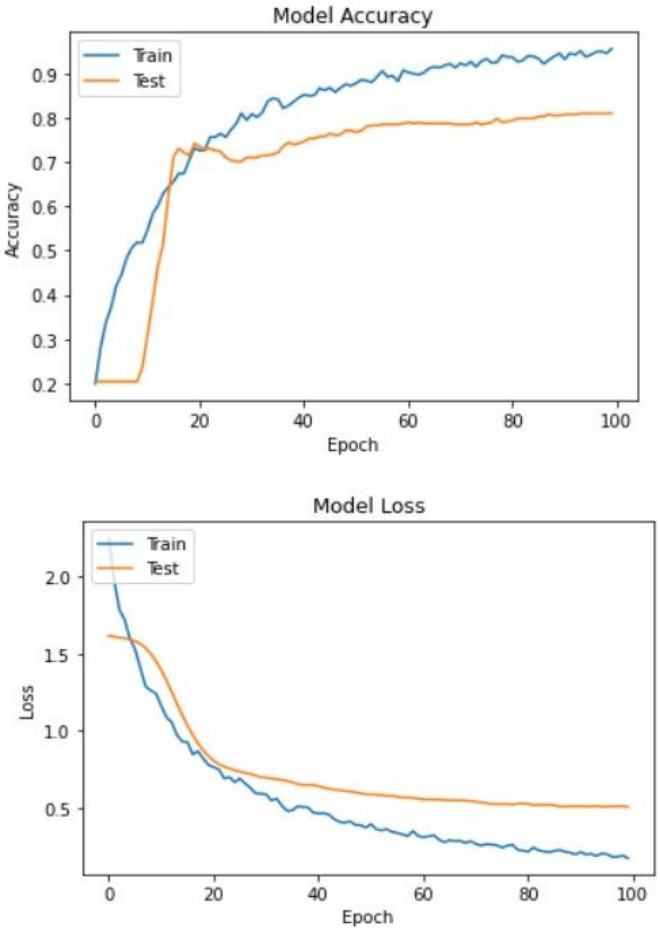
Fig. 5. This figure defines the sequence diagram of the project. It has been giving a brief information about the training and testing datasets. It specifies the input and results that are obtained from the dataset implementation.

6. Result

The problem our project was trying to solve and help our users with was resolved up to a greater extent as the accuracy and sensitivity we received was greater than the once achieved by the previous attempts by various people. Here we tested our data on various algorithmic model to decide the best one suited for the job which in our case provided us with the accuracy of around 98%.

Conclusion We started a search to provide the solution to the people for improving their posture's accuracy and to make and environment where they learn and improve at the comfort of their home even if they do not have a physical trainer or if they can't afford one. And to our hardwork and luck we were able to create a platform to where they can see how correct they are in the process and even improve themselves as our model provides you with an accuracy of your posture to you for your reference helping you in correcting it and improving yourself with an easy to use user interface.

Classifier	Accuracy
Ridge Classifier	98.12
Random Forest Classifier	97.37
Gradient Boosting Classifier	98.87
K-Neighbors Classifier	93.25



7. Future Scope

With the blazing speed at which computer vision is evolving, new pose estimation techniques and models will soon replace the tried-and-true methods of today. Using better models and techniques in the future for improvement of the model is something we'd certainly be interested in. Extending our work with a number of other popular yoga poses is also on the cards. Improving the model, we develop with new, improved and thorough datasets with a higher usability could also be possible. Adding a whole new design language to the project by introducing UI elements is also possible. Developing a web application or a native app will also help in the distribution of this idea.

8. References

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

1. 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

2. A. Jalal, N. Sarif, J. T. Kim, and T.-S. Kim, "Human activity recognition via recognized body parts of human depth silhouettes for residents monitoring services at smart home," Indoor and Built Environment, vol. 22, no. 1, pp. 271–279, 2013.View at: Publisher Site | Google Scholar

3. Y. Zhan and T. Kuroda, "Wearable sensor-based human activity recognition from environmental background sounds," Journal of Ambient Intelligence and Humanized Computing, vol. 5, no. 1, pp. 77–89, 2014.View at: Publisher Site | Google Scholar

4. H. T. Chen, Y. Z. He, C. L. Chou, S. Y. Lee, B. S. P. Lin, and J. Y. Yu, "Computer-assisted self-training system for sports exercise using kinects," in Proceedings of the 2013 IEEE International Conference on Multimedia and Expo Workshops (ICMEW), pp. 1–4, IEEE, London, UK, July 2013.View at: Google Scholar

5. A. Mohanty, A. Ahmed, T. Goswami, A. Das, P. Vaishnavi, and R. R. Sahay, "Robust pose recognition using deep learning," in Proceedings of the International Conference on Computer Vision and Image Processing, pp. 93–105, Springer, Singapore, December 2017.View at: Google Scholar 10