

Deep learning-based yoga posture specification using OpenCV and media pipe

Chunduru Anilkumar^{1,3}, Robbi Jyothsna¹, Sattaru Vandana Sree¹, E. Gothai²

¹Dept of Information Technology, GMR Institute of Technology, Rajam, Andhra Pradesh-532127

²Department of Computer Science and Engineering, Kongu Engineering College, Erode, India

³anilkumar.ch@gmr.it.edu.in

Abstract. Yoga is a year's discipline that calls for physical postures, mental focus, and deep breathing. Yoga practice can enhance stamina, power, serenity, flexibility, and well - being. Yoga is currently a well-liked type of exercise worldwide. The foundation of yoga is good posture. Even though yoga offers many health advantages, poor posture can lead to issues including muscle sprains and pains. People have become more interested in working online than in person during the last few years. People who are accustomed to internet life and find it difficult to find the time to visit yoga studios benefit from our strategy. Using the web cameras in our system, the model categorizes the yoga poses, and the image is used as input. However, the media pipe library first skeletonizes that image. Utilizing a variety of deep learning models, the input obtained from the yoga postures is improved to improve the asana. On non-skeleton photos, VGG16, InceptionV3, NASNetMobile, YogaConvo2d, and also InceptionResNetV2 came in the order of highest validation accuracy. The proposed model YogaConvo2d with skeletal pictures, which is followed by VGG16, reports validation accuracy in contrast, NASNetMobile, InceptionV3, and InceptionResNetV2.

Keywords: deep learning, media pipe, classification, skeletonization, posture detection.

1. Introduction

People typically assume that yoga could be a kind of exercise that has stretching and folding of the piece however Yoga is far quite a simple exercise. Yoga could be an approach to life or the Art of living through the mental, religious and physical path. It permits us to attain stillness and to faucet into the consciousness of the inner self. It additionally helps in learning a way to rise on top of the pull of mind, emotions, and lower bodily wants and face the challenges of day-to-day life. Yoga works on the amount of one's body, mind, and energy. Regular observation of yoga brings positive changes within the practice – robust muscles, flexibility, patience, and physiological state. even as moving into the yoga cause properly is vital, therefore starting up is the correct approach. get laid with awareness, coordination every body movement with the breath as you gently begin the cause and enter the resting position Human cause estimation has returned to a protracted approach within the last 5 years, astonishingly hasn't surfaced in several applications simply however. this is often a result of additional focus has been placed on creating cause models larger and additional correct, instead of doing the engineering work to form

them quickly and deploy-able all over. With movement, our mission was to style and optimize a model that leverages the most effective aspects of progressive architectures, while keeping illation times as low as doable. The result's a model which will deliver correct key points across a good kind of poses, environments, and hardware setups. Using a depth camera makes it possible to detect disguised hands and facial expressions as well as the entire articulation of mortal hands. Utilizing a portable moving depth camera, low-cost, three-dimensional (3D) reconstruction of terrain has been investigated. Reason of the sterility essential item in the operating room, touch-free, interfaces getting popular in medical applications. Algorithms for estimation and shadowing of mortal disguises are usually learned from sizable pool training data. Viscosity assessment of mortal disguise data reveals the data online observance structure, which can then be used for shadowing and disguise estimation. Compared to the data used to estimate the viscosity of mortal disguise, our stuff is different. Performance in reading is impacted by biased training data. Using a weight retrogression to deleting these training datasets. There are numerical methods for altering data distribution from established data viscosity.

2. Related work

Presented the YOGI data-set, which has roughly 5500 photos. Using the "tf-posture," Different angles and points from these visual frames were fed into a number of machine learning models, such as Logistic Regression, followed by Random Forest, Support Vector Machine (SVM), Decision Tree, followed by Nave Bayes, and also K Nearest Neighbour (KNN). The accuracy of the given algorithm random-forest classifier was 99.04%. [1].

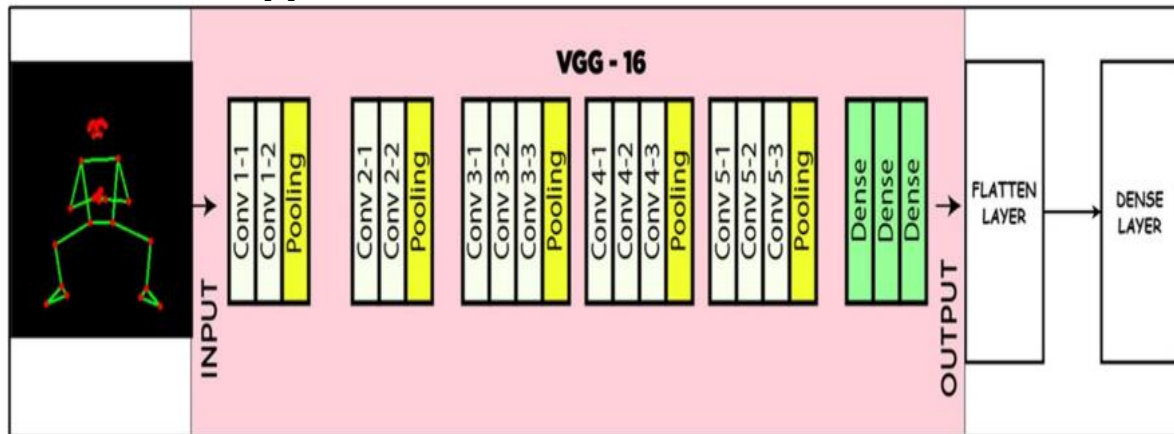


Figure 1. Architecture of VGG-16.

In order to get beyond the drawbacks of being state-of-the-art, they proposed a vision-grounded system in this work for the real-time recognition of yoga activities. To do this, we first create a sizable data-set of ten Yoga positions that were photographed in challenging real-world settings so that the built system might operate more effectively when placed in those settings. Twenty-seven people of all ages participated in the creation of the internal data set of yoga poses (8 males and 19 ladies). The videos, which feature acts in all their conceivable variants, were recorded using MI Max and also One Plus 5T smartphones in both indoor and outdoor settings. Second, we suggested a lightweight 3D CNN armature that detects Yoga positions by taking use of the crucial spatial-temporal link between them. Yoga disguise sequences can be investigated to improve the planned system's subtlety in recognition. For the real-time recognition of yoga poses, the developed method can be be improved and implemented on moveable bedded bias.[2].

A key area of computer vision has been how the human body responds to visuals. The estimate problem for mortal disguise is made significantly easier by the use of a depth camera. created a body component identification algorithm for the commercial Kinect device. The recognition performance was improved by using tentative retrogression timbers to incorporate knowledge and global variables, as the

stoner's height and also branch lengths. A data-driven strategy specifically keeps database of mortal acts and seeks the best-matching acts at runtime to grease disguise rebuilding has been studied solely. [3].

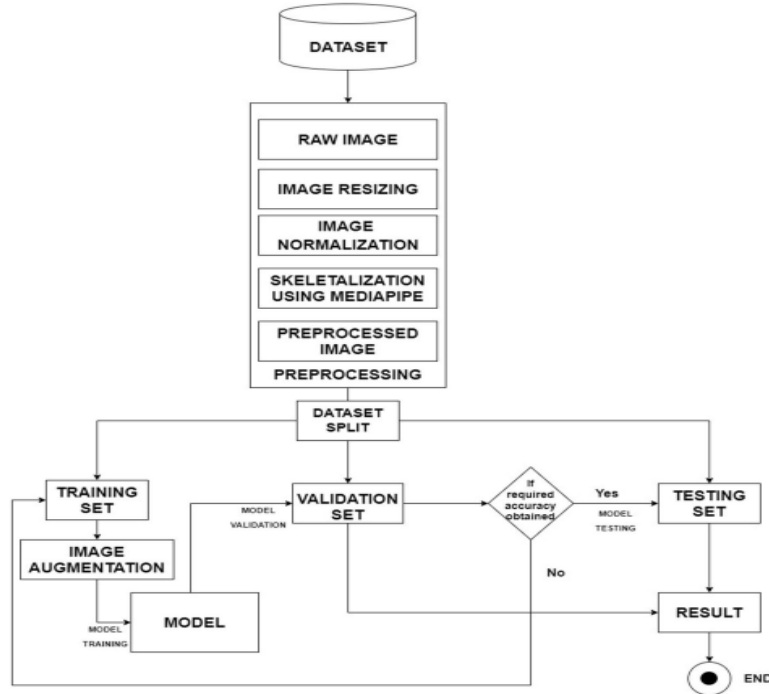


Figure 2. Flowchart for pose recognition.

A yoga recognition system that makes use of an ordinary RGB camera. The data-set is intimately available and was gathered using HD 1080p Logitech webcams on 15 people (10 men and 5 women). To photograph the stoner and describe important details, Open Pose is employed. End-to-end deep literacy-grounded frames do not require handmade features, therefore new asanas may be added by simply again training t model with fresh data. For purpose of discovering yoga postures, the method of utilizing both CNN and LSTM on data obtained from Open Pose has been designed to be largely successful. For 12 people, the system can identify the six asanas both in real-time and on recorded films (five males and seven ladies). Real-time testing and data collection have been done by many people. Yoga acts on a videotape are successfully detected by the system with 99.04 frame-wise delicacy and 99.38 frame delicacy after polling 45 frames. The system successfully produced 98.92 dishes in real-time.[4].

Both ANN and FCM classifier were trained using 30% of the subjects in this study, while the remaining 70% were used for testing. Three posture cases were differently chosen from each type of posture among all subjects during the training phase. Each data frame was fed into the associated ANN classifier algorithm to determine the classification order and corresponding FCM classifier algorithm to determine the identify outcome. The average degree of delicacy for posture detection was 89.34%, and it ranged from 70% to 100%. The present data frame was quickly acknowledged as the first posture. The ultimate recognition outcome in our recognition styles was calculated using accretive probability. As a result, several instances position three was recognized as posture one, which decreased the identification delicateness. [5]

With the introduction of Kinect, a computer-assisted self-training system has been developed. With the extraction from the skeleton for pose recognition that adopted three postures: downward-facing dog stance, warrior 3 posture, and tree posture, the devices typically contain RGB cameras, infrared projectors, and detectors. During the experiments, an overall accuracy of 82.84% was attained. [6]

OPENPOSE is a real-time multi-person human posture recognition library that, for the first time, identifies important facial and body components in a single photo. With the aid of underlying layers, OpenPose has been used to recognize the introduction's major parts and extract specifics from the input image.[7]

For self-training, a yoga posture recognition method using the Kinect camera has been created. 300 films of 12 yoga poses being performing five times each by five different yoga practitioners were collected. Once the body has been extracted from the video clip, a skeleton is used to depict the yoga positions. Produced accuracy was 99.33%. [8]

Table 1. Comparison table.

Author	Title	Techniques	Advantages
Shubham Garg et.al (2022)	Yoga pose classification: Both CNN and also Media Pipe inspired deep learning approach for real-world application	CNN, Media pipe, Classification	The obtained results by both media pipe and non-media pipe versions are examined to see if skeletonization of the yoga pose image increases the model performance or not.
Shrajal jain et.al (2020)	Three-dimensional CNN-inspired deep learning architecture for yoga pose recognition in the real-world environment	Convolution Neural networks, Conv 3D and max pooling (Maxpool3D)	In this paper, poses are captured by a camera as well as recorded videos of the poses feeding the sequence to the proposed the 3D CNN neural network model, which was developed, had 99.96% of training accuracy and 99.49% of validation accuracy.
Kyungyong Yang et.al (2015)	Controllable data sampling in the space of human pose	Bayesian network model, PCA techniques	Three categories are created from the subsampled motion capture data. There were 8 699 Type I postures, 7 300 Type II poses, and 2 386 Type III positions as a consequence of the classification.
Santosh Kumar Yadav et.al (2019)	Real-time Yoga recognition using deep learning	CNN, LSTM	Yoga posture recognition is quite good. For 12 people, the system can identify the six asanas both in real time and on recorded films (five males and seven females). Real-time testing and data collection have been done by many people.
Jiwen Zhang Et.al (2019)	Yoga Posture Recognition and Quantitative Evaluation with Wearable Sensors Based on Two-Stage Classifier and Prior Bayesian Network	ANN classifier, FCM classifier (Futures commission merchant)	Bayesian network is trained by the database. Subject two and followed by five performed many standards given to the evaluation results by the Bayesian network while gaining on, FCM classifier algorithm membership computed is very high

Table 1. Continue.

Author	Title	Techniques	Advantages
Chhaihyuo Long et.al (2021)	Development of a yoga posture coaching system using an interactive display based on transfer learning	CNN, Logistic Regression, SVM, KNN, naive Bayes	Consideration of 12 poses that are captured through webcams. The poses are skeletonized and application of feature extraction and classification through Machine Learning models with 96% of accuracy.
Deepak Kumar et.al (2020)	Yoga Pose Detection and Classification Using Deep Learning	CNN, RNN, MLP, SVM	By providing the input, the model detects the accurate pose among the learned 6 poses.
Munkhjargal Gochoo et.al (2019)	Novel IoT-Based Privacy-Preserving Yoga Posture Recognition System Using Low-Resolution Infrared Sensors and Deep Learning	DCNN, CNN, Adaboost algorithm,	Trained with 300 clips of 12 yoga poses through the help of a Kinect camera (a depth cam that sees in 3D and creates a skeleton image and detects the movements). The body contour is extracted first and then skeletonized and produces 99.33% of accuracy.

Table 1 explains the comparison over the different papers along with their advantages, title name and respective authors.

3. Methodology

In this model, the main factor, to fete the asana, is the angle and distance between the joints. Using Media pipe, milestones of different body corridors & joints will be detected and also arc tan (function of NumPy) will calculate the angles between those milestones. According to the angles, it'll estimate the posture and also recoup the instructions of that asana and it'll read out the instructions. An overall architecture of the given system is shown in illustration below.

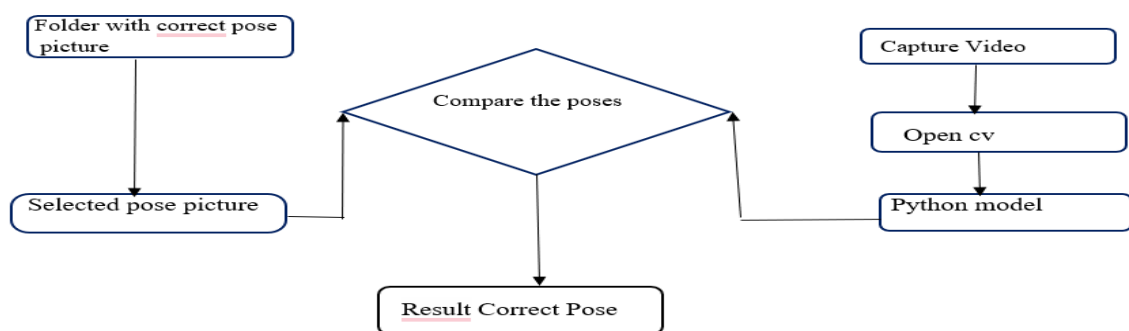


Figure 3. Flowchart of the procedure.

The visual is gains by a camera, it be divides camera module on a smartphone, now extensively accessible, or a webcam, it is useful fashion to capture prints because nearly everybody has one of these types of captured input results. The system's input element is the camera. A webcam, a mobile camera, or a divides camera module can be used as the source.

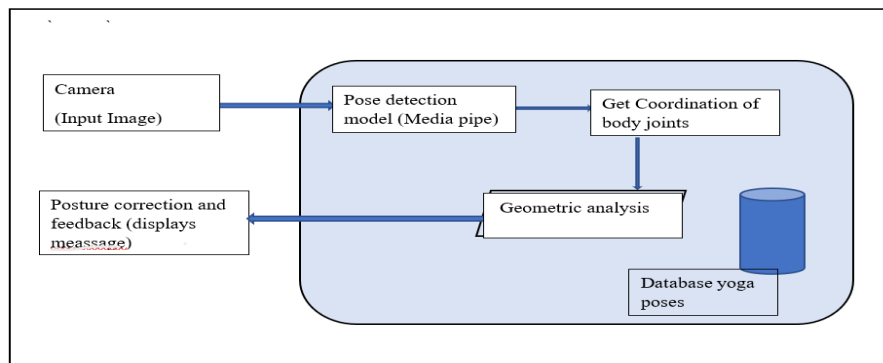


Figure 4. Block diagram.

The camera is used to receiving images and give total data to the model After receiving the visualized item, we developed a model using CNN[9][10][11]. The proposed system can fete a wide range of positions. As a conclusion, exercising data sets to the topmost extent doable. Pose identification is done using Media Pipe the stoner's input data is taken. Produce a correct shell exposure of the stoner using this data. Milestones on the mortal body gestures identify the crucial points and places

4. Conclusion

We have created a system in this design that consists of a channel for disguise identification, point localization on the mortal body, and an error identification method. This technique attempts to aid individuals in correctly yoga practice on their own and assist with ailments that may result from improper yoga poses. The approach is suitable for evaluating the stoner's disguise from the front and providing feedback so that they can improve their yoga disguise using deep literacy techniques. The designed model is mounted atop a dashboard that was likewise made with stoners in mind.

References

- [1] Garg, Shubham Saxena, Aman Gupta, Richa 2022 2022/06/03 Yoga pose classification: a CNNand MediaPipe inspired deep learning approach for real-world application Journal of Ambient Intelligence and Humanized Computing <https://doi.org/10.1007/s12652-022-03910-0>
- [2] Jain, S., Rustagi, A., Saurav, S., Saini, R., & Singh, S. (2020). Three-dimensional CNN-inspired deep learning architecture for Yoga pose recognition in the real-world environment. Neural Computing and Applications. doi:10.1007/s00521-019-04232-7
- [3] Yang, K., Youn, K., Lee, K., & Lee, J. (2015). Controllable data sampling in the space of human poses. Computer Animation and Virtual Worlds, 26(3-4), 457–467. doi:10.1002/cav.1662
- [4] Yadav, S. K., Singh, A., Gupta, A., & Raheja, J. L. (2019). Real-time Yoga recognition using deep learning. Neural Computing and Applications. doi:10.1007/s00521-019-04232-7
- [5] Wu Z, Zhang J, Chen K, Fu C. Yoga Posture Recognition and Quantitative Evaluation with Wearable Sensors Based on Two-Stage Classifier and Prior Bayesian Network. Sensors (Basel). 2019 Nov 23;19(23):5129. DOI: 10.3390/s19235129. PMID: 31771131; PMCID: PMC6929085.
- [6] Long, Chhaihuoy Jo, Eunhye Nam, Yunyoung 2022 2022/03/01 Development of a yoga posture coaching system using an interactive display based on transfer learning The Journal of Supercomputing 5262 EP - 5284 <https://doi.org/10.1007/s11227-021-04076-w>
- [7] D. Kumar and A. Sinha, "Yoga Pose Detection and Classification Using Deep Learning," International Journal of Scientific Research in Computer Science, Engineering and

- Information Technology, pp. 160–184, Nov. 2020, DOI: 10.32628/cseit206623.
- [8] M. Gochoo et al., "Novel IoT-Based Privacy-Preserving Yoga Posture Recognition System Using Low-Resolution Infrared Sensors and Deep Learning," in *IEEE Internet of Things Journal*, vol. 6, no. 4, pp. 7192-7200, Aug. 2019, DOI: 10.1109/JIOT.2019.2915095.
 - [9] Easwaramoorthy, S., Sophia, F., & Prathik, A. (2016, February). Biometric Authentication using finger nails. In *2016 international conference on emerging trends in engineering, technology and science (ICETETS)* (pp. 1-6). IEEE.
 - [10] Easwaramoorthy, S., Moorthy, U., Kumar, C. A., Bhushan, S. B., & Sadagopan, V. (2018). Content based image retrieval with enhanced privacy in cloud using apache spark. In *Data Science Analytics and Applications: First International Conference, DaSAA 2017, Chennai, India, January 4-6, 2017, Revised Selected Papers 1* (pp. 114-128). Springer Singapore.
 - [11] Subramanian, M., Cho, J., Sathishkumar, V. E., & Naren, O. R. (2023). Multiple types of Cancer classification using CT/MRI images based on Learning without Forgetting powered Deep Learning Models. *IEEE Access*.