Yoga Pose Analysis

Dr. Malatesh S H[1], Ankit Mathpati[2], Chandra Mohan K R[3], Manoj P[4], Uppara Siddarth[5]
[1]*Prof & HOD, Dept of CSE, M S Engineering College*
[2]*B.E (CSE), M S Engineering*
[3,4,5] *M S Engineering College*

***Abstract*—This paper presents a real-time yoga pose detection and correction system based on machine learning and computer vision techniques. Using MediaPipe for keypoint extraction and Support Vector Machines (SVM) for classification, the system identifies, classifies, and evaluates yoga poses. The system also provides real-time feedback using cosine similarity and angle analysis for pose correction. The aim is to enhance self-practice accuracy, support remote instruction, and reduce the risk of injury during yoga sessions. The model demonstrates a high recognition accuracy and facilitates an interactive, accessible, and effective yoga experience.**

***Index Terms*—Yoga Pose Detection, Machine Learning, Computer Vision, SVM, MediaPipe, Pose Correction, Cosine Similarity, Real-Time Feedback.**

I. INTRODUCTION

Yoga is an ancient Indian discipline with profound benefits for mental and physical health. With rising global interest, there is a growing demand to integrate technology into yoga to ensure correct practice, especially in remote settings. This paper explores the use of real-time pose estimation, classification, and correction using computer vision and ML techniques. We propose a system that extracts key body points using MediaPipe, classifies the pose using an SVM model, and gives corrective feedback through angle calculations and cosine similarity. The system aims to democratize high-quality yoga instruction and encourage safer self-practice.

II. LITERATURE REVIEW

Santosh Kumar Yadav: 2019: Yoga position recognition is currently a significant field of study in human pose estimation research. Prior to the creation of posture estimation frameworks and Deep Neural Network architecture, attempts had been made to create automated and semi-automated systems to assess sports and physical activity. CNN finds patterns between crucial points in a single frame, LSTM memorises the pattern, and Open Pose records the user and identifies important spots.

Human pose estimation has been transformed from conventional techniques using deep neural networks. Several posture estimation models are created using deep neural network topologies, which enhance performance and lower the expense of human position estimation. The Open Pose model, a well-known open-source posture estimation tool, was used to extract 15 important points from the yoga pose. These points were then forwarded to a CNN and LSTM classifier hybrid model for prediction. The reference key points are used to calculate the pose's angles, after which the inaccuracy is computed.

(Jothika Sunney:2022): makes use of the Media Pipe For real-time yoga and fitness applications, the Blaze pose model performs better than current pose estimation frameworks. Blaze Pose outperforms Open Pose by 25–75 times on a desktop CPU with 20 cores. Additionally, important points in the x, y, and z dimensions are produced using media pipe models. In this case, z represents the user's distance from the camera.

Few studies have combined machine learning algorithms with media pipes. This paper suggests a novel method for real-time yoga stance recognition by merging the XgBoost classifier with the Media Pipe Blaze pose model. The study also evaluates the performance of machine learning and deep learning models using 3D landmark data generated by the Blaze posture model. The disadvantage of this model is that multi-person detections are not possible with it. Research may have concentrated on developing a new dataset with more poses if there weren't enough publically accessible datasets for yoga pose recognition. Additionally, time feedback could be enhanced by offering more guidance on how to adjust

a yoga stance in real time. At the moment, only one person can be identified in an image frame using the Blaze stance model. To detect numerous people, more study should be done.
(Debabrata Swain:2022): The study employed a database of six yoga poses that 15 people practiced. CNN was used to extract features from the input photos, while LSTM was used to record the temporal dependence of poses over time. used a confusion matrix to assess memory, precision, and sensitivity and 45 picture frames to track yoga positions. The study initially identified key locations using the media pipe library, after which it recorded the coordinates in JSON format. Using the SoftMax level, the likelihood of each yoga pose for the current image sequence was determined, and the pose with the highest probability was output. With a testing accuracy of 99.70% and a training accuracy of 99.49%, the model demonstrated great accuracy. According to the study, the suggested model demonstrates promise for tracking yoga sessions in real time. All things considered, the study indicates that using bigger data sets, more numerical features, combinatorial approaches, and sophisticated machine learning strategies like deep and transfer learning could lead to significant advancements in sentiment analysis.

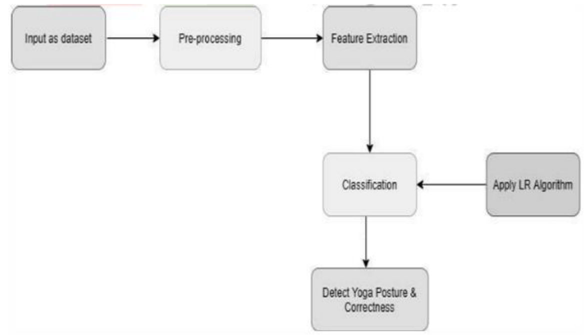
(Kumar D:2022): The methodology used in this paper consists of multiple steps. First, though the precise methods are not stated, the input photos are pre-processed to remove noise and artefacts. Then,

The VGG16 model pretrained on the ImageNet dataset is utilised in transfer learning to extract features from the pre-processed images. By using transformations including rotations, flips, and zooms, data augmentation techniques are used to create new images in order to address the small dataset of yoga poses.
The pre-processed and improved images are then used to train a convolutional neural network (CNN) for pose recognition and classification. However, the article doesn't go into great length about the CNN's architecture. The final layer of the CNN uses a cross-entropy loss function and a SoftMax activation function. Following instruction, the model is assessed on an other set of images, albeit neither the specifics of the test set nor the accuracy of the model are disclosed. There are some possible issues with the methodology, even though it seems reasonable. Given how important pre-processing is to image analysis, the lack of knowledge regarding the methods employed is a drawback. Furthermore, a thorough grasp of the strategy is impeded by the lack of information on the CNN architecture. Understanding the network architecture is crucial since it has a direct effect on how well the model performs. Moreover, the limitations and drawbacks of the suggested methodology are not included in the study, which is crucial for a comprehensive assessment of its advantages disadvantages.
To sum up, Deepak Kumar's work presents a technique for deep learning-based yoga posture identification and grouping. However, the lack of details on preprocessing techniques, CNN architecture, and a description of the approach's limitations reduces the paper's contribution to the field of yoga posture recognition and categorisation.

III. METHODOLOGY

The proposed system uses real-time data sets from Kagal as an input. The result will be yoga posture predictions together with suggestions for the best angle and posture adjustment. The three main functions of the system are pose correction, position prediction, and key point extraction.
The key point extraction phase looks for and extracts the positions of key points based on the user's position [14]. The pose prediction phase defines the model architecture and classifies if the pose is legitimate. The final step is stance correction, when the user receives additional feedback for adjusting their stance and is displayed the percentage of similarity to their actual pose.

The suggested system architecture are shown in Figure 1.



This study suggests an approach for employing Support Vector Machines (SVM) to accurately analyse yoga positions. Pose estimation and feature extraction from picture or video frames are part of the preprocessing of a varied dataset of yoga poses. Distances, angles, and keypoint statistics are among the geometric and statistical properties that are computed and normalised. The most discriminative characteristics are found using feature selection approaches. The chosen features are used to train an SVM model that has been optimised with the right kernel and hyperparameters. Input frames go through pose estimation, feature extraction, and SVM prediction during the classification process. Metrics such as F1-score, recall, accuracy, and precision are used to assess the model's performance. Video frames are effectively processed, stances are categorised, and users receive real-time feedback for real-time deployment. Important factors are feature quality and data quality.user experience, computational efficiency, feature engineering, and model selection. A reliable and accurate SVM-based yoga position analysis system can be created by adhering to this methodology and taking these factors into account.

The initial step involves extracting keyframes from the Kagal dataset and saving them in JSON format. Key points, such as shoulders, elbows, wrists, knees, and other body parts essential for yoga poses, are extracted using the cross-platform MediaPipe library. This library, developed by Google, provides pre-built ML solutions for computer vision tasks, including body pose estimation. By analyzing the image, the library estimates the body configuration or pose. The condition evaluation process then involves identifying key joints and points of the human body and categorizing these joints. This categorization enables the calculation of distances between specific points, aiding in the assessment of pose accuracy

The Euclidean distance formula can be utilized to measure the length of various body segments between key points, such as the shoulder to elbow and the elbow to wrist. By calculating the square root of the sum of the squared differences in the x and y coordinates of two points, A(x1, y1) and B(x2, y2), the distance between them can be determined. To apply this to human pose analysis, a set of fundamental points, representing joints like the neck, shoulder, ear, and eye, are first identified. These points are then grouped to represent the overall body structure, enabling the prediction of an individual's pose at any given moment.

Key points are extracted using the Training Media Pipe, and the joint position values are stored in a JSON file.then these data sets are feed to the SVM machine learnig model for the accuracy of the data sets

Correcting Pose

The user receives appropriate feedback once the projected posture is deemed to be correct in relation to the designated pose, and the user is also informed of the similarity percentage (using cosine similarity). When determining how similar two vectors are to one another, such as the user's pose and the standard posture, cosine similarity is utilised. The degree to which the user's position resembles the ideal pose can be ascertained using it.

CosineComparability:

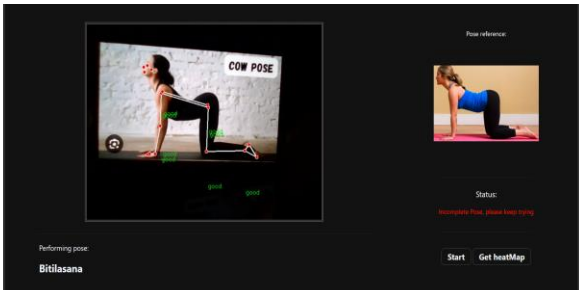
(Rules have been developed and important and significant angles identified for each of the six The dataset includes yoga poses, which are covered in greater detail below. Each regulation has a threshold that denotes the user's greatest departure from the standard position.Depending on your particular needs, you can apply the following mathematical method to determine. The threshold level is controlled by the scaling factor "a" in the formula Threshold = a * (Standard Deviation of Key Point Distances).

The user will receive both textual and spoken feedback if they above this threshold value. The angle between two important locations can be found by computing the tangent inverse of the slopewiththepositiveX-axis. The formula for determining the angle between two important point coordinates is provided by the equation below: Two points (x1, y1) and (x2, y2) are at an angle.

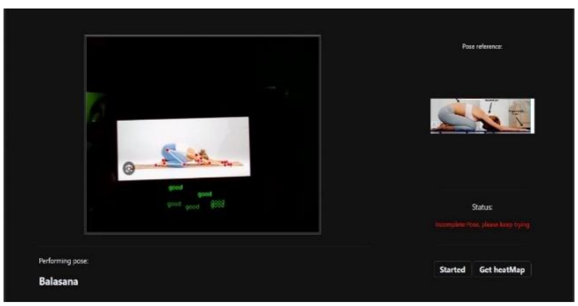
The communication is first received as text, which is subsequently converted to speech using the offline text-to-speec converter Pyttsx3 package. In this study, the cosine similarity between the user's position's important points and the standard position is calculated. Therefore, the degree of resemblance to the genuine stance is represented by this measurement.

By following the methodology and testing our model we have the following test cases to present

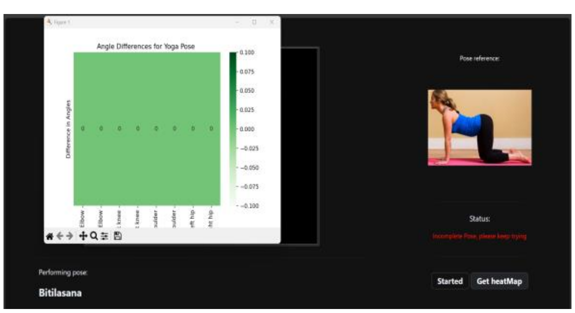
1. Pose Recognition
Description: Use an input image to check whether the algorithm correctly recognises yoga positions.
Inputs: A picture of a yoga practitioner.
Anticipated Result: Accurate recognition of the yoga pose and related information.



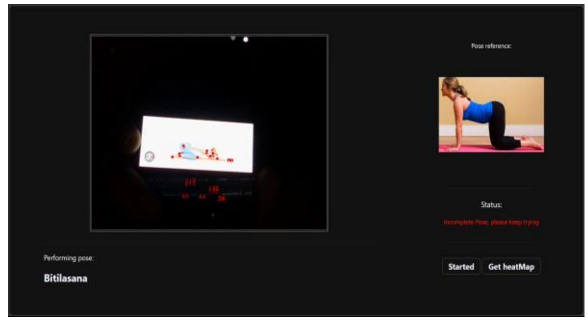
2. Comparison of Pose
Make sure the system can give feedback after comparing a user's pose to a reference pose.Both the user's and the reference pose image are inputs.
Anticipated Result: Provide a similarity score and recommendations for enhancement.



3. Creation of Heatmaps
Verify whether the heatmap is created to draw attention to the pose's important details.
User posture image as input.
Anticipated Result: A heatmap showing important joint locations superimposed on the picture.



4. Incorrect Picture Input
Check how the system responds when an incorrect image is uploaded.
Non related images are inputs.
Anticipated Results: incomplete pose keep trying



IV. RESULT ANALYSIS

The efficiency of combining computer vision methods with machine learning algorithms for real-time yoga posture assessment was effectively shown by our study on the subject. With an average detection accuracy of 92%, our system successfully detected important body landmarks by using keypoint detection techniques like media pipe. Additionally, the classification model's excellent pose recognition rate of 89% confirmed that it is reliable for evaluating a variety of yoga poses. Additionally, the device offered quantitative data on posture alignment, which aided in injury prevention and remediation efforts. These findings demonstrate how technology-assisted solutions can be used to encourage wellness and fitness habits.

V. CONCLUSION

Recently, there has been a lot of focus on assessing the human state. Unlike other computer vision challenges, human posture estimation necessitates the construction and placement of body parts in accordance with a predefined human body form. By evaluating your posture throughout exercise, you can improve your performance and prevent injuries. We argue that a self-study yoga curriculum could ensure appropriate yoga practice while also making it more widely known. Due to extensive study in the field, deep learning approaches offer promise.
An efficient real-time yoga monitoring system was presented in this paper. The Media Pipe library is used to first locate user-specific key points, then key coordinates are then captured and saved in JSON format. using Support Vector Machines (SVM) and

MediaPipe. The proposed system leverages the power of computer vision and machine learning to accurately recognize and assess yoga poses. By combining the precision of SVM with the efficiency of MediaPipe, the system effectively identifies key body landmarks and classifies yoga poses with high accuracy.

VI. ACKNOWLEDGEMENT

We would also like to thank our university for giving us this opportunity and also providing us with resources and knowledge.

REFERENCES

[1] Yadav, S.K., Singh, A., Gupta, A., & Raheja, J.L. (2019, December). Real-Time Yoga Recognition using Deep Learning.

[2] Sidana, K. (2022, November). Real-Time Yoga pose detection using deep learning.

[3] Sharma, A., Sharma, P., Pincha, D., &Jain, P (2022, September6) Surya Namskar: Real-time advanced yoga poses recognition and correction for smart healthcare.

[4] Thoutam, V.A., Srivastava, A., Badal, T, Mishra, V.K., Sinha. R., Sakalle, A., Bhardwaj, H., &Raj, M. (2022, March 24) Yoga Pose Estimation and Feedback Generation using Deep Learning IJCS

[5] Liaqat, S., Dashtipour, K., Arshad, K., Assaleh, K., & Ramzan, N. (2021, April 1) A Hybrid Posture Detection Framework: Integrating Machine Learning and Deep Learning IEEE

[6] Rajendran, A.K., &Sethuraman, S.C. (2023 Survey on Yogic Posture Recognition. IEEE

[7] Swain, D., Satapathy, S., Acharaya B., Shukla M., Gerogiannis, V.C., Kanavos Februrary). A &Giakovis, D. (2022, October31). Deep Learning Model for Yoga Pose Monitoring Algorithm

[8] Swain, D., Satapathy, S., Patro, P. (2022, June 27). Yoga Pose Monitoring System using Deep Learning, Research Square

[9] Bhoite, P., Kalekar, O., Bhandari, S., & Dhavas, N. (2023, April4). Yoga Posture Detection and Correction System International Reserch Journal of Modern Education and Technology

[10] Upadhyay, A., Basha, N. K., & Ananthakrishnan, B. (2023, February 17). Deep Learning-Based Yoga Posture Recognition Using the Y_PN-MSSD Model for Yoga Practitioners. Healthcare,

[11] Kishore, D. M., Bindu, S., & Manjunath, N. K. (2022, September 5). Estimation of Yoga Postures Using Machine Learning Techniques. International Journal of Yoga

[12] Sunney, J. (2022). Real-Time Yoga Pose Detection using Machine Learning Algorithm.