Educational Administration: Theory and Practice
2024, 30(5), 14735-14741
ISSN: 2148-2403
https://kuey.net/
Research Article

XAI-Driven Yoga Pose Analysis and Correction in Real Time

Kruti Shah[1*], Manasvi Gupta[2], Nilesh Patil[3], Sridhar Iyer[4], Fatema Dolaria[5], Rishita Koshtiz[6]

[1*,2,3,4,5,6]Computer Engineering, DJ Sanghvi College Of Engineering, Mumbai, India
[1]krutishah957@gmail.com, [2]manasvig476@gmail.com, [3]nilesh.p@djscce.ac.in, [4]sridhar.iyer@djsce.ac.in, [5]fdolaria@gmail.com

Citation: Kruti Shah et al (2024) XAI-Driven Yoga Pose Analysis and Correction in Real Time, *Educational Administration: Theory and Practice*, *30*(5), 14735-14741
Doi: 10.53555/kuey.v30i5.7384

ARTICLEINFO **ABSTRACT**

Abstract—Yoga is an integral part of one's physical and mental well-being.Improper postures could reduce its benefits or cause injury. This study introduces a real-time yoga pose detection and correction system using Explainable Artificial Intelligence (XAI). It utilizes an advanced pose estimation model to precisely identify key body landmarks and analyzes the precision of yoga poses performed by an user in real-time. XAI techniques provide transparent and interpretable feedback so that the user can perceive and correct in real time any kind of misalignment. The integration of XAI ensures not only improved accuracy of poses but also empowers the practitioner with instant actionable insights toward a safer and effective practice. Such a system is well adapted for deployment within personalized virtual yoga training to realize real-time guidance on improving overall wellbeing.

Index Terms—Shap, XAI, Yoga, Random Forest Classifier, Feedback, Pose detection, Error Threshhold, MediaPipe,Pose Correction.

I. INTRODUCTION

An ancient practice, yoga combines physical postures,mediation and breath control. It has gained global recognition for its physical and mental health benefits.It originated over 5,000 years ago in ancient India, yoga began as a spiritual discipline detailed in the Vedas and Upanishads. The classical period saw the Yoga Sutras of Patanjali, which systematized yoga into an eight-limbed path aimed at achieving mental discipline and spiritual liberation. In the 19th and 20th centuries, globalization of yoga occurred and yoga spread to the Western world and adapted it to modern practice. Yoga today is taking a new look, providing a blend of the ancient teaching along with the modern approach towards physical health, mental clarity, and spiritual well-being.

As more and more individuals take up yoga, it is very important to perform various yoga postures or asanas correctly to derive maximum benefits from them and prevent injuries. Traditional approaches of learning yoga involve in-person instruction, where teachers provide guidance and corrections to their students. With the rise of digital platforms, there is a increasing need for automated systems that can accurately detect and correct yoga poses in real-time.

Recent advancements in the field of machine learning have paved the way for developing such systems, enabling the detection of human poses from images or video streams with remarkable accuracy.However, the complexity and opacity of such models often make it difficult for humans to fully understand and interpret decisions, creating a "black-box problem". This is very important when the output directly influences users' physical practices, such as in the correction of yoga poses.

The present research investigates the applicability of XAI techniques for yoga pose identification and correction. The model used is Random Forest, which is a very robust classifier, and thus, it helps in making decisions more interpretable. XAI is used to render the developed system more transparent, hence rendering the user the capability of actually understanding the underlying rationale behind the feedback.

This paper presents a comprehensive approach to developing and evaluating an XAI-based yoga pose detection and correction system. The system combines the advanced pose estimation capabilities of Mediapipe along with the predictive power of the Random Forest classifier model and enriching the system with XAI techniques.The proposed solution not only detects and corrects yoga poses but also provides interpretable feedback on pose accuracy. The proposed system would help every kind of practitioner, from novices to advanced, to be able to achieve the proper posture and, thus ensures increased safety and efficiency when practicing yoga.

II. LITERATURE REVIEW

The literature review presented in the paper explores the existing tools, systems, algorithms and methodologies in Realtime Yoga Pose Detection and Correction using XAI, addressing the need for fitness monitoring systems for individuals.By examining diverse research projects and studies, it aims to identify strengths, limitations and challenges in the solutions. This review sets the groundwork for proposing a comprehensive solution to bridge existing gaps.

In their work Saurav, et al.[1] evaluated several architectures including a hybrid CNN and LSTM model and three 3d CNN models.The 3D CNN model achieved an accuracy of 99.65% .They achieved an 8 FPS for the 3D CNN model thereby enhancing the computational efficiency and making the system feasible for real-world applications.They use transfer learning with pre-trained models in order to improve both accuracy and speed.Despite the improvements they have certain limitations .The model lacks overall transparency and explainability for corrections of the yoga poses, which may hinder the user's understanding of the feedback and their ability to apply it effectively.

Chaudhari, et al.[4]implemented a system based on deep learning for correction of yoga poses in real-time along with identification of poses using a Convolutional Neural Network by tracking 15 key body points.The system attains 95% accuracy in the identification of the postures.Despite its effectiveness, the system has notable limitations.The pose identification model struggles with reduced accuracy for poses that are similar to each other.

The implementation by Astuti, et al.[2] evaluates the accuracy of yoga poses using pose estimation with ResNet and cosine similarity for the comparisions of positions. COCO dataset is used to train the model and attains a precison of87% and recall of 88.2%.The application exhibits several limitations, primarily related to the variability in model performance across different yoga poses. It predicts certain poses with a high accuracy and performs significantly lower for other poses.The model's effectiveness is highly-pose dependant and thus may not generalise for a broad range of yoga poses.It limits the utility for comprehensive yoga practice evaluation.

Thar, et al.[3] elaborated on the development of a Yoga posture assessment method using OpenPose.The system was tested on three individuals having distinct body shapes, genders and ages.Also three different yoga poses were used. The system compares angles between some body joints of a user and an instructor, and if they exceed a threshold, it suggests corrections.The results are indicated by a gradient color system: green for correct alignment and red for large discrepancies.

Lobo, et al.[5]explained the development of a system to analyze and correct yoga poses in real time.It captures video input, processes frames to identify landmarks on the body and calculates the joint angles.They are compared with prerecorded correct angles stored in a database.When discrepancies are found, feedback is given by the system to adjust the posture. The results indicate that, although the system is adequate at tracking and correcting poses fairly well does not allow for individual variations in anatomy.The application provides real-time feedback on yoga poses but there is no transparency in its assessments.

Dittakavi, et al.[6] discussed a system for recognizing and correcting poses majorly focusing on home fitness scenarios.A Densenet classifier was employed along with a K-Nearest Neighbors(KNN) to enhance the interpretability and performance.They performed experiments on Yoga-82, Pilates-32 and Kungfu-7 datasets which demonstrated that the DenseNet+KNN outperformed the KNN models upto 18% for noisy,self-occluded poses, offering superior explanations by highlighting key joints.The limitations include lack of explainablity of models.

In conclusion, the literature review underlines both advances and limitations in Real-time Yoga Pose Detection and Correction using XAI.The systems though accurate, function as black-box systems providing little insight into rationale behind their decisions.

II. PROPOSED METHODOLOGY

XAI is important for the transparency and trustworthiness of the system in yoga pose detection and correction. XAI gives ways in which practitioners and developers understand how the model reaches its decisions. It gives insight into the features or landmarks that most significantly influence pose assessment. This is critically important to make sure that the recommendations provided by the model are accurate and reliable, especially in a domain like yoga, where wrong form can lead to injury. Also, XAI makes validation and refinements of the model quite easy for the practitioner, so the performance of the model increases with time when new poses come up in yoga, which makes it really safe to perform those poses."

A. Dataset Used

The model is trained using the Yoga-82 dataset [8] which contains 28,478 images representing 82 different classes of yoga poses. The dataset provided a robust foundation for training the Random Forest model in effective classification of diverse yoga poses while striking a balance between computational efficiency and model performance. This dataset contains a wide range of poses, from basic to advanced, capturing various angles, body orientations, and levels of difficulty.

B. *Pose Extraction*

The pose extraction process starts by capturing a video frame in real-time and determining key body points using MediaPipe.The system identifies 33 landmarks of the body such as shoulders, elbows, ankles and feet.The detected landmarks are represented as x, y, z coordinates along with a visibility score.After this, the landmarks are re-arranged into a DataFrame.

C. *Pose Identification*

Once the extraction is completed, the 33 key landmarks of the user's body are then analyzed by a trained machine learning model Random Forest classifier to identify the yoga pose.The coordinates are then analyzed to match the pose against a set of known yoga postures in the dataset.The Random Forest Classifier processes the coordinates to predict the specific yoga pose based on how the body's key points align with those in the model's training data.The system then evaluates the confidence of the prediction and if it meets a predefined threshold, it displays the name of the pose.

D. *Error Estimation and Feedback*

The system uses the angle formula given in Eq.(1) to calculate the angle between three points, such as the shoulder, elbow and wrist in 2D space.Eq.(2) converts the angle from radians to degrees.This calculated angle helps to determine the alignment of body parts, which is crucial when correcting yoga poses. The yoga pose detection system evaluates errors by comparing the measured joint angles to predefined target values that are obtained from the model.The system allows minor deviations within a ±20-degree range . The system identifies a misalignment if the joint deviates from the the angle more than the specified threshold.

$$\text{angle} = \arctan2(c_y - b_y, c_x - b_x) - \arctan2(a_y - b_y, a_x - b_x) \tag{1}$$

$$\text{angle}_{degrees} = \text{angle}^{radians} \times \frac{180}{\pi} \tag{2}$$

Based on these deviations, feedback is provided to the users offering guidance on how to adjust their joint positions for improved alignment.Say, if a user's joint angle deviates from the required, users are instructed to either move the joint closer or further away from the body.The feedback generated for the user includes the degree of adjustment required thus ensuring precise corrections of the yoga poses.The process is also visualized through a bar chart as shown in Fig.(1) that shows the percentage difference between the measured and target angles which helps users understand and correct their posture effectively.For accessibility, the system incorporates text-to-speech (TTS) functionality to verbally announce the identified pose and the feedback generated, assisting users who may have visual impairments or disabilities. The feedback is shown and communicated in real-time.

E. Feedback using XAI

The system developed integrates Explainable AI (XAI) techniques, specifically SHAP SHapley Additive exPlanations) values, to provide transparent and actionable feedback.The model starts by predicting a yoga pose based on features such

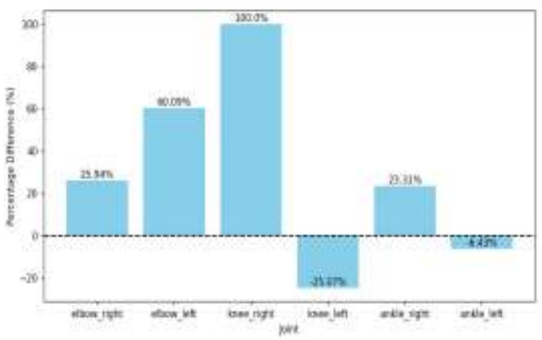


Fig. 1. Percentage error difference for each joint of a pose.

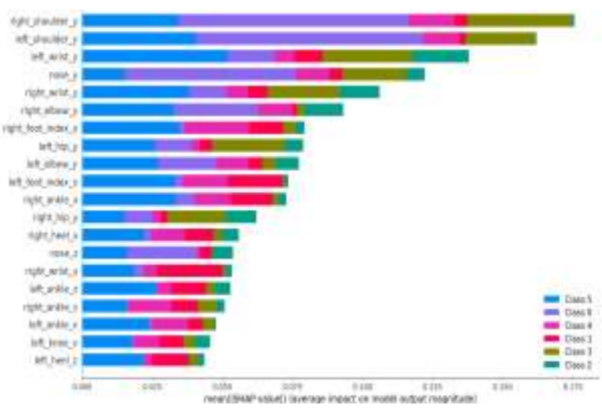


Fig. 2. Impact of each feature on model

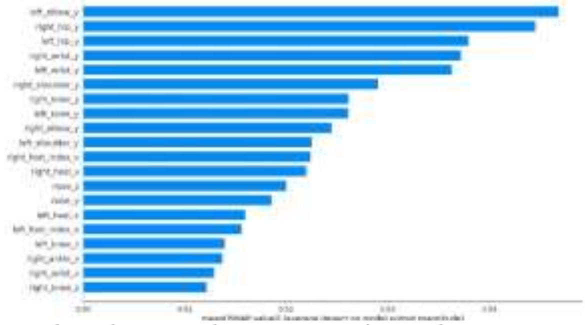


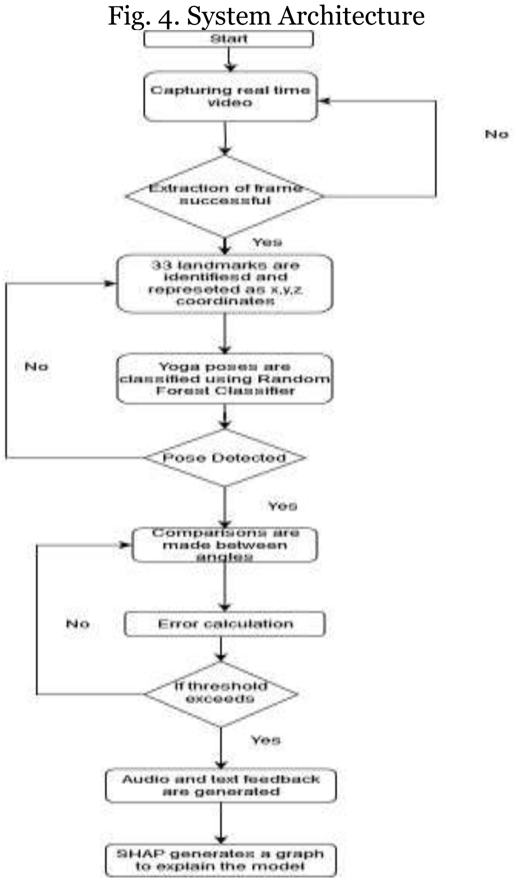
Fig. 3. Landmarks contributing most for Padmasana pose

as joint angles and landmarks.SHAP values are then calculated in order to determine how each feature contributes in the prediction.The chart in Fig.(2) visualises the SHAP values by highlighting the contribution of each feature in the model's decision making process.
SHAP values are also utilized in order to explain the model's predictions by identifying which particular joint angles or landmarks contributed the most to pose classification decision.This has been illustrated in Fig(3).for the Padmasana pose.The detailed feedback obtained helps users understand why certain corrections are necessary making the advice more targeted and actionable.

IV. SYSTEM OVERVIEW

The proposed system recognizes six yoga asanas in real time .These asanas include Downdog, Vrikshasana, Goddess, Virabhadrasana2, Phalakasana and Padmasana. The flowchart in Fig.(4) demonstrates the proposed system architecture.
Step 1: Start by capturing user's real-time video when he is performing yoga poses. Then, image extraction is applied in order to extract all the body landmarks.
Step 2: 33 landmarks are that are extracted are represented as x, y and z coordinates along with a visibility score. The image is then classified using a Random Forest Classifier to detect the yoga pose being performed.
Step 3: If the pose is correctly identified, the system compares it with the expert image from the dataset and calculate the differences in the joint angles.
Step 4: f the angle deviations exceed a set threshold, audio and text feedback are provided to the user. Additionally, using SHAP algorithm the model's decision is explained.

Fig. 4. System Architecture



Thresholding Mechanism Error

The system includes a critical decision point which uses an error thresholding approach hence feedback is only generated when errors from the ideal pose are large enough to justify a correction. The user's pose is compared with a reference image then using this comparison, an error metric is computed and compared with a pre-defined threshold which is set to 0.85.In case the error is below this threshold, then the system does not provide the feedback, which ensures that small deviations which do not affect the efficacy and safety of the pose are not unnecessarily interrupted. The mechanism enables the system to focus only on meaningful corrections, making the user's experience as best as possible by giving guidance when it is really necessary.

A. *Integration of SHAP for Interpretability*

The final step in the pipeline which incorporates SHAP algorithm is a perfect illustration of this system's commitment to transparency and empowerment of users. By creating clear, detailed graphs, the system not only gives feedback on how to improve but also explains why these suggestions are being made.This approach explains how the model's recommendations are affected by a variety of factors such as poses and landmarks thus turning the black box into something interpretable and building trust in the model. The users gain more insight into processes of decision making, enabling them to correct their poses effectively and simultaneously learn the underlying principles of good practice in yoga for both physical and mental well-being.

B. *Real Time Constraints and Optimization*

Though the system is structured to perform in real time in an effective manner ,it can be optimised to achieve faster processing times and more efficient resource usage. Future work includes leveraging hardware acceleration techniques to be beneficial in resource-constrained environments such as mobile devices where processing power and memory are limited.These optimizations will bring a system that is more responsive for the user, real-time with less powerful hardware setups, and open to a much broader audience with wider accessibility and usability.

V. RESULTS AND DISCUSSIONS

TABLE I PERFORMANCE METRICS

Asana Names	precision	recall	f1-score	support
Downdog	0.97	0.97	0.97	49
Goddess	0.92	0.94	0.95	32
Padmasana	0.93	0.95	0.93	23
Phalakasana	0.88	0.88	0.88	28
Virabhadrasana-ii	0.93	0.96	0.99	28
Vriksasana	0.91	0.89	0.98	38
accuracy	0.97			198
macro avg	0.92	0.93	0.95	198
weighted avg	0.95	0.94	0.95	198

This table is evaluating the performance of a model to classify yoga poses. Precision, recall, and F1-score are listed for each pose to show how the model identifies correctly posed individuals and notates correct identifications without false positives or negatives. It also shows overall accuracy, macro average, and weighted average of all the poses. Higher values in these metrics signify better classification accuracy. The accuracy score is 0.97.

Fig. 5. Feedback generated for Padmasana

The image in Fig.5 presents a visual analysis of a person's Padmasana, posture. A digital overlay of lines and points is superimposed on the individual to provide a framework for assessment. Key areas of the body, such as the knees, ankles, and elbows, are indicated with specific feedback. For instance, the image might highlight that the knees are aligned correctly for the pose while suggesting an adjustment to the angle of the right elbow for optimal alignment and comfort. This type of feedback can be instrumental in helping practitioners refine their form and prevent potential injuries.

Fig. 6. Feedback generated for Goddess

The image in Fig.6 provides feedback on a person’s Goddess pose. The person is attempting the pose, and the image includes a diagram overlaying the person with lines and points to analyze their posture. The image also provides textual feedback for specific body parts, highlighting areas that need improvement (right elbow angle, left elbow angle, right knee, right ankle) and areas that are correct (left knee, left ankle).

VI. CONCLUSION

In this project a system is developed to aid people in their practice of proper and safe yoga, hence reducing injuries caused by poor techniques. The system uses an advanced pose detection pipeline to detect 33 key body landmarks.

These landmarks are represented as x, y, and z coordinates and there is a visibility score associated with the detected points.Once the landmarks are found, the error identification process occurs through a comparison of the user’s pose with the dataset’s pose.Thus using this comparison, the system is able to properly identify deviations and give personalized feedback. The feedback is provided as both audio and text formats, directing the user on how to correct their pose to achieve better alignment and technique. This enables the user to self-correct their pose and significantly improve their yoga sessions by fostering the independence in their practice.This approach makes yoga sessions better, and at the same time, the users themselves are able to self-correct and fostering Independence in their practice.

For increased interpretability of the feedback, SHAP(SHapley Additive explanations) algorithm is employed. SHAP visualizes the contribution of each feature in the model, which gives the user an insight into the factors that influenced the system analysis and suggestions.This level of transparency will help users understand the reason behind feedback.

Though this system is able to categorize and analyze the yoga poses, some factors come into play in determining the accuracy of pose detection with the proposed system, such as good lighting, occlusion if a limb goes out of view, or nonstandard postures. In these cases, the challenges can lead to less fine-tuned landmark detection, which makes the feedback it provides probably less reliable.

REFERENCES

1. Saurav, S., Gidde, P., & Singh, S. (2024). Exploration of deep learning architectures for real-time yoga pose recognition. Multimedia Tools and Applications, 1-43.
2. Astuti, A. D., Karlita, T.,& Asmara, R. (2023). Yoga Pose Rating using Pose Estimation and Cosine Similarity. Jurnal Ilmu Komputer dan Informasi, 16(2), 115-124.
3. Thar, M. C., Winn, K. Z. N., & Funabiki, N. (2019, November). A proposal of yoga pose assessment method using pose detection for self-learning. In 2019 International conference on advanced information technologies (ICAIT) (pp. 137-142). IEEE.
4. Chaudhari, A., Dalvi, O., Ramade, O., & Ambawade, D. (2021, June). Yog-guru: Real-time yoga pose correction system using deep learning methods. In 2021 international conference on communication information and computing technology (ICCICT) (pp. 1-6). IEEE.
5. Lobo, A., Lobo, J., Joy, A., & Mirza, I. A. (2022, August). Yoga correction using machine learning. In 2022 2nd Asian Conference on Innovation in Technology (ASIANCON) (pp. 1-6). IEEE.
6. Dittakavi, B., Bavikadi, D., Desai, S. V., Chakraborty, S., Reddy, N., Balasubramanian, V. N., ... & Sharma, A. (2022). Pose tutor: an explainable system for pose correction in the wild. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 3540-3549).
7. Ashraf, F. B., Islam, M. U., Kabir, M. R., & Uddin, J. (2023). Yonet: A neural network for yoga pose classification. SN Computer Science, 4(2), 198.

8. Verma, M., Kumawat, S., Nakashima, Y., & Raman, S. (2020). Yoga82: 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).
9. Hsu, Y. L., Yang, S. C., Chang, H. C., & Lai, H. C. (2018). Human daily and sport activity recognition using a wearable inertial sensor network. IEEE Access, 6, 31715-31728.
10. Singh, A., Agarwal, S., Nagrath, P., Saxena, A., & Thakur, N. (2019, February). Human pose estimation using convolutional neural networks. In 2019 amity international conference on artificial intelligence (AICAI) (pp. 946-952). IEEE.
11. Yadav, S. K., Agarwal, A., Kumar, A., Tiwari, K., Pandey, H. M., & Akbar, S. A. (2022). YogNet: A two-stream network for realtime multiperson yoga action recognition and posture correction. KnowledgeBased Systems, 250, 109097.
12. Sharma, A., Shah, Y., Agrawal, Y., & Jain, P. (2022). iYogacare: realtime Yoga recognition and self-correction for smart healthcare. IEEE Consumer electronics magazine, 12(4), 47-52.
13. Upadhyay, A., Basha, N. K., & Ananthakrishnan, B. (2023, February). Deep learning-based yoga posture recognition using the Y _PN-MSSD model for yoga practitioners. In *Healthcare* (Vol. 11, No. 4, p. 609). MDPI.
14. Wei, G., Zhou, H., Zhang, L., & Wang, J. (2023). Spatial–Temporal SelfAttention Enhanced Graph Convolutional Networks for Fitness Yoga Action Recognition. Sensors, 23(10), 4741
15. Vallabhaneni, N., & Prabhavathy, P. (2023). Segmentation quality assessment network-based object detection and optimized CNN with transfer learning for yoga pose classification for health care. Soft Computing, 1-23.
16. Hsu, C. C., Morsalin, S. S., Reyad, M. F.,& Shakib, N. (2023, December). Artificial Intelligence Model Interpreting Tools: SHAP, LIME, and Anchor Implementation in CNN Model for Hand Gestures Recognition. In International Conference on Technologies and Applications of Artificial Intelligence (pp. 16-29). Singapore: Springer Nature Singapore.
17. Speith, T., Crook, B., Mann, S., Schomacker, A., & Langer, M.¨ (2024). Conceptualizing understanding in explainable artificial intelligence (XAI): an abilities-based approach. Ethics and Information Technology, 26(2), 40.