



A DEEP LEARNING-BASED REAL-TIME YOGA ASSISTANCE FOR THE VISUALLY IMPAIRED

Ms.REMYA PV

*Assistant Professor CSE Department
St Thomas College of Engineering and Technology
Kannur,India*

Fathima PK

*Computer Science Engineering
St Thomas College of Engineering and Technology
Kannur,India*

Fathimathul Misba Shareen VP

*Computer Science Engineering
St Thomas College of Engineering and Technology
Kannur,India*

Fathima Zuhara PK

*Computer Science Engineering
St Thomas College of Engineering and Technology
Kannur,India*

Shahana Shirin CP

*Computer Science Engineering
St Thomas College of Engineering and Technology
Kannur,India*

Abstract—Yoga offers many physical, mental, and emotional benefits, and having a trained instructor by our side can significantly enhance the experience and outcomes. Here, we've suggested a method that increases yoga's accessibility and inclusivity for everybody, particularly the blind. Users can personalize their yoga experience, get real-time feedback, and even request particular exercises. The voice assistant communicates through speech-to-text and text-to-speech technologies. Using a deep-learning Convolutional Neural Network (CNN), the system is trained to predict and correct yoga positions based on key body locations. The suggested yoga posture estimate system takes advantage of the OpenPose method in MediaPipe, a popular computer vision technique for identifying and following key spots on the human body in pictures or movies. OpenPose, Mediapipe, SpeechRecognition, and CNN are index terms.

Index Terms—OpenPose technique, Mediapipe, SpeechRecognition, CNN.

I. INTRODUCTION

In this study, we present a novel approach that ensures the accessibility of yoga for those with visual impairments. Accessing the traditional approaches is limited to them, which rely on visual clues and human instructors. Our system provides autonomous assistance and guidance to all kinds of people despite their liabilities, especially the visually impaired, using voice-based assistance. A deep learning Convolutional Neural Network (CNN) combined with the OpenPose estimation using mediapipe is the basis of our system. This

combination makes it possible to precisely estimate and adjust yoga postures according to important body spots, guaranteeing a customized and successful practice. With the use of simple voice commands, tailored feedback that is provided instantly, and requests for specific routines, users may engage with the system with ease. The work has brought innovation in technology and artificial intelligence thus it is indispensable.

II. MOTIVATION

Yoga has many health benefits emotionally, physically, and mentally. Integrating it into our daily lives is inevitable for all the benefits it provides. Performing yoga does need a trainer or guidance. Performing yoga without proper guidance or knowledge would result in stress and strains on the body. Having a yoga assistant or trainer is not that easy for everyone. Having one at their disposal would bring about complications. People tend to access recorded videos of yoga poses online but it does not provide any guidance. Having easy-to-access and simple technologies would encourage the users to anticipate yoga into their daily lives therefore, they enhance their wellbeing. When considering well-being every kind of person has to be considered like the visually impaired. Integration of natural language processing is needed to provide benefits to them, that is it should be accessible to every kind of users despite of their liabilities.

III. LITERATURE REVIEW

A. Machine Learning-Based Yoga Pose Identification Using TF-Pose Estimation

This paper presents a novel approach for precise identification of yoga poses leveraging machine learning techniques. A meticulously curated dataset, named YOGI, comprising 5500 high-resolution images capturing ten distinct yoga poses, was utilized. The tf-pose estimation algorithm was employed for real-time generation of human body skeletons, facilitating feature extraction of joint angles. Through rigorous experimentation with six machine learning models, including Logistic Regression, Random Forest, SVM, Decision Tree, Naive Bayes, and KNN, the Random Forest Classifier exhibited exceptional performance with 99.04% accuracy. Preprocessing steps ensured uniformity, while feature extraction focused on 12 crucial angles for pose detection. This comprehensive methodology achieved an impressive overall accuracy of 94.28%, underscoring the transformative potential of machine learning in advancing yoga posture recognition. The findings contribute to both scientific understanding and practical applications, promising technology-enhanced yoga practices.

B. MTMD: Mixup Triplet Learning for Mahalanobis Distance in Open-Set Human Activity Recognition

This paper introduces MTMD, a novel approach addressing open-set Human Activity Recognition (HAR) challenges by leveraging mixup triplet learning. HAR from wearable sensor data often encounters dynamic signals and undefined activities over time, necessitating robust open-set classifiers. MTMD tackles this by enhancing discriminative features using mixup triplets, effectively distinguishing between known and unknown activities. The method establishes an embedding space in the training phase using mixup triplets and triplet loss, followed by an inference phase employing Mahalanobis distances for open-set recognition. Experimental results on benchmark datasets showcase MTMD's superior performance, outperforming existing methods across various metrics. MTMD offers a pioneering training methodology with potential for future research in deep metric learning and incremental learning in HAR. MTMD introduces a robust distance-based open-set recognition approach using Mahalanobis distances and thresholds derived from the learned embeddings. By calculating distances between test embeddings and representative embeddings for each known class, MTMD effectively distinguishes between known and unknown activities, enhancing the reliability of HAR systems in dynamic environments. This innovative approach not only advances the field of sensor-based HAR but also holds promise for applications in diverse real-world scenarios where accurate activity recognition is crucial for human well-being and safety.

IV. METHODOLOGY

A. Deep Learning Model:

A symmetric approach leveraging MediaPipe for real-time pose detection was used to create the deep learning model, a

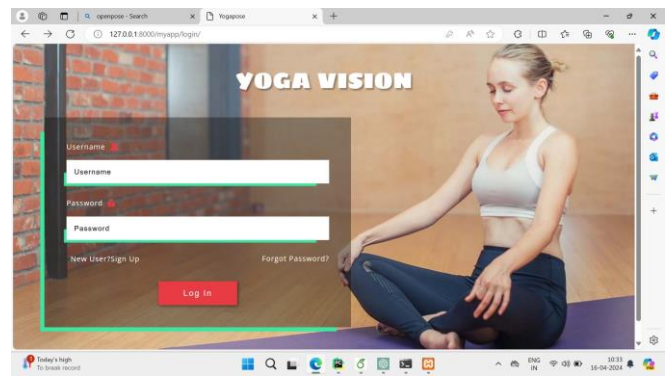


Fig. 1. Login Page

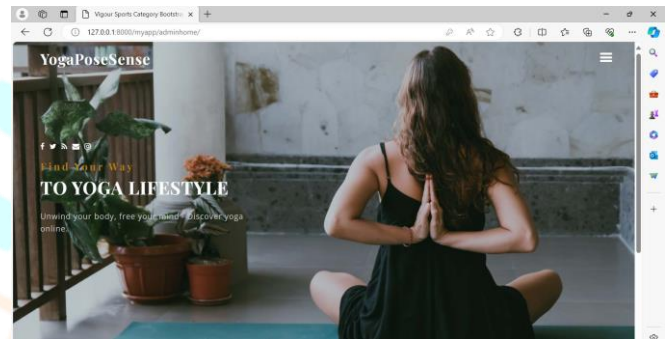


Fig. 2. Home Page

Neural Network (NN) trained using Keras for pose recognition, and OpenPose was used for real-time pose recognition. The deep learning model was created by training a CNN in Keras with a dataset comprising various distinct yoga poses along with their labels. Yoga poses are collected through data collection, which is done in real time and these poses are utilized to train the neural network model. This trained model is integrated into the real-time pose recognition algorithm.

B. Website

The website was developed with a combination of HTML, CSS, and Python Django for the front end. As DeepLearning and Python have a strong integration, the backend code was written in Python. The website is designed so that it provides an easy to use environment to the user.

V. IMPLEMENTATION

The data will be collected in real-time using the open cv library. The open-pose estimation technique will analyze the frames and store the relative coordinates of each pose landmark concerning the first landmark thus making use of the MediaPipe framework. The openpose technique is implemented in MediaPipe as part of the pose detection pipeline. The landmark data of each pose will be saved as a NumPy array. The data collection will be done until it takes 80 frames. The NumPy data would be further processed and trained. The training will be done by the CNN model with three dense layers: two hidden layers with 128 and 64 neurons,

using the hyperbolic tangent activation function, and an output layer with softmax activation for multi-class classification. We compile the model using the RMSprop optimizer and categorical cross-entropy loss function. The trained model will then be saved to the system. Keras is the primary deep learning framework for designing, compiling, and training the neural network model. The pytsx3 library is used for text-to-speech conversion, and the speech recognition library's recognizer object is initialized and used to capture audio input from the microphone. Afterward, the recorded audio is processed through the recognize Google method to identify speech using Google's voice recognition service. These libraries improve the system by giving users an interactive voice-based experience.

VI. WORKING

The proposed system is designed to process real-time visuals of yoga performances that are captured through a webcam to detect the yoga poses. The initial stage involves a Pre-processing step, which is responsible for refining and optimizing the input visuals. The openpose detection algorithm works in real-time using Media Pipe, a tool that facilitates the extraction of key points and landmarks from the webcam video frames. Following the preprocessing step, the visuals are given to the OpenPose detection method. The OpenPose method extracts essential features from the yoga performance visuals, it also identifies the key points and joint positions that are crucial for understanding the body posture during different yoga poses. the output from the OpenPose method, containing the feature information, is then directed to a Sequential Convolutional Neural Network (CNN). This component is involved in both the training and testing phases.

During training, the CNN learns to recognize and categorize different yoga poses based on the extracted features. Once the CNN has undergone training, it moves into the testing phase, where it evaluates its performance on new and unseen yoga performance visuals. After the testing phase, if any inaccuracies or errors are detected in the recognized yoga poses, the system provides voice commands for corrective actions. This suggests a user-friendly interface where users can verbally interact with the system to improve and adjust the identified poses, enhancing the overall user experience. To facilitate efficient data management, both for training and testing purposes, a database is integrated into the architecture. This approach combines advanced algorithms with traditional yoga to create an immersive platform for the users.

After the testing phase, if any inaccuracies or errors are detected in the recognized yoga poses, the system employs voice commands for corrective actions. This suggests a user-friendly interface where users can verbally instruct the system to improve and adjust the identified poses, enhancing the overall user experience. To facilitate efficient data management, both for training and testing purposes, a database is integrated into the architecture. This approach combines advanced algorithms with traditional yoga, creating an immersive platform for refining poses. Users can easily improve their practice, fostering a deeper connection with yoga.

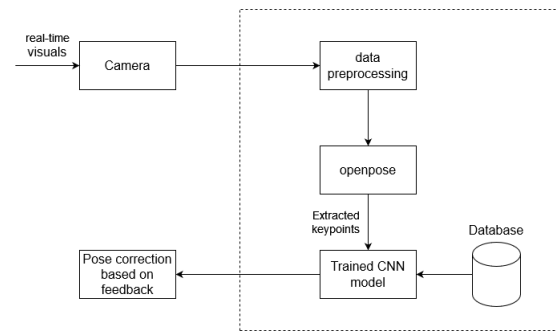


Fig. 3. Architecture Diagram



Fig. 4. Body Keypoints

VII. RESULTS

The system recognizes a total of five yoga pose. The system takes frames when the full body is visible. For each pose it checks for errors and corrections. When the yoga pose is incorrect it reads out the description of how to perform the correct yoga pose. Based on the commands from the user the system moves to the next pose or exits the applications. All five yoga poses were recognized accurately and corrections were given in case of any errors.

The five poses that are recognized by the system are mountain pose, Urdhvahastasana, Tree pose, Warrior pose, and Lotus pose. The poses get checked in the same order given here. On receiving the command "help" the system gives out the tip for the current yoga pose in the iteration. On receiving the command "next" it moves to the next yoga. And for the command "exit" the application exits by stopping the pose estimation model.

VIII. CONCLUSION

The real-time yoga assistance system, tailored for individuals, particularly those with visual impairments, leverages

advanced technologies like OpenPose detection in MediaPipe and Convolutional Neural Networks (CNNs). By processing visual data during yoga sessions, it identifies and tracks key body points to provide personalized guidance. Through real-time feedback conveyed via voice prompts, users can adjust their poses for improved practice. Additionally, the system allows for interactive engagement through voice commands, enhancing accessibility and inclusivity in wellness activities. Overall, it represents a significant step towards making yoga more accessible and enjoyable for all individuals, emphasizing inclusivity and empowerment in health and wellness.

IX. FUTURE SCOPE

The future scope of our system can accommodate multiple users simultaneously would enable group yoga sessions. Which will track and analyze the movements of multiple individuals in real-time. Also combining the human yoga instructors with AI-driven technology could offer users a personalized learning experience. Incorporating VR technology could offer users immersive and interactive yoga experiences, allowing them to practice in virtual environments tailored to their preferences. A dataset with diverse yoga poses performed by individuals of varying body types, ages, and skill levels would improve its accuracy and adaptability.

X. ACKNOWLEDGEMENT

We extend our appreciation to St. Thomas college of engineering and Technology for their support in the development of this proposed system. Special thanks to Mr. Jithin S for their guidance and expertise throughout the project. We also acknowledge the contributions of all team members who dedicated their time and efforts to research, design, and implement the system. Additionally, we are grateful to the participants who provided valuable feedback and insights during the testing phase.

REFERENCES

- [1] Chaudhary, I., Singh, N.T., Chaudhary, M. and Yadav, K., 2023, May. Real-Time Yoga Pose Detection Using OpenCV and MediaPipe. In 2023 4th International Conference for Emerging Technology (INCET) (pp. 1-5). IEEE.
- [2] Anuradha, M., Rao, A., Shreyas, S. and Sanjaya, K.C., 2023, April. Real Time Virtual Yoga Tutor. In 2023 IEEE 8th International Conference for Convergence in Technology (I2CT) (pp. 1-4). IEEE.
- [3] Lee, Minjung, and Seoung Bum Kim. "Sensor-Based Open-Set Human Activity Recognition Using Representation Learning With Mixup Triplets." IEEE Access 10 (2022): 119333-119344.
- [4] Sharma, A., Agrawal, Y., Shah, Y. and Jain, P., 2022. iYogacare: real-time Yoga recognition and self-correction for smart healthcare. IEEE Consumer Electronics Magazine.
- [5] Anwarul, S. and Mohan, M., 2022, October. Deep Learning-based Yoga Pose Recognition System using Hyperparameter Tuning. In 2022 10th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions)(ICRITO) (pp. 1-6). IEEE.
- [6] Tarek, O., Magdy, O. and Atia, A., 2021, December. Yoga Trainer for Beginners Via Machine Learning. In 2021 9th International Japan-Africa Conference on Electronics, Communications, and Computations (JAC-ECC) (pp. 75-78). IEEE.
- [7] Gupta, A. and Jangid, A., 2021, September. Yoga pose detection and validation. In 2021 International Symposium of Asian Control Association on Intelligent Robotics and Industrial Automation (IRIA) (pp. 319-324). IEEE.

- [8] Chaudhari, A., Dalvi, O., Ramade, O. and 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.
- [9] Rishan, F., De Silva, B., Alawathugoda, S., Nijabdeen, S., Rupasinghe, L. and Liyanapathirana, C., 2020, December. Infinity yoga tutor: Yoga posture detection and correction system. In 2020 5th International conference on information technology research (ICITR) (pp. 1-6). IEEE.
- [10] Agrawal, Y., Shah, Y. and Sharma, A., 2020, April. Implementation of machine learning technique for identification of yoga poses. In 2020 IEEE 9th international conference on communication systems and network technologies (CSNT) (pp. 40-43). Ieee..



International Research Journal
IJNRD
Through Innovation