Real Time Yoga Asana Recognition using Deep Learning

[1]M Darshan, [2]Appu V, [3]Rohan M, [4]Nandan N, [5]Nandita Bangera

1,2,3.4 Student .5Assistant Professor

Computer Science and Engineering

RV Institute of Technology and Management,Bengaluru,VTU

Abstract Real-time recognition of yoga postures is an important task in the field of computer vision and human-computer interaction. In this article, we propose a deep neural network architecture for real-time detection of yoga postures. The proposed architecture is based on a combination of Convolutional Neural Networks (CNN) and Long-Term Memory Networks (LSTM). CNN networks are used to extract features from the input image, while the LSTM network is used to model the temporal dynamics of yoga poses. The proposed model was trained on a large data set of yoga poses and achieved a "high accuracy of 96".3% on a 10-piece test set..

Keywords - Real-time yoga pose recognition, Deep Neural Networks , Convolutional Neural Networks, long short-term memory networks, temporal dynamics, computer vision, human-computer interaction.

I. INTRODUCTION

Yoga, as a form of exercise, encompasses a variety of body positions and movements. Engaging in regular yoga practice offers numerous health benefits, including enhanced flexibility, strength, and relaxation. Nevertheless, beginners often struggle with mastering proper posture and alignment in yoga. To enrich the yoga experience and provide precise feedback, there is a growing interest in the development of computer vision systems capable of recognizing and tracking yoga postures. Deep learning has emerged as a powerful tool in computer vision, enabling accurate object recognition in images and videos. In this study, we present a novel camera-based system that utilizes deep learning algorithms to recognize and track yoga poses. Our system leverages convolutional neural networks (CNN) to identify body parts and employs localization algorithms to determine body position and orientation. To estimate the overall posture, we utilize a clustering method to group these body parts. The proposed system boasts several potential applications, such as automated yoga instructions, real-time feedback, and personalized exercise recommendations. For instance, it can guide beginners through step-by-step instructions for correctly performing various yoga poses. Additionally, the system can monitor practitioners' progress over time and offer personalized recommendations based on their performance. This paper outlines the design and implementation of our camera-based system, evaluates its performance using a dataset of yoga videos, compares it with existing yoga posture detection methods, and discusses its advantages and limitations. Our findings collectively underscore the potential of deep learning approaches to accurately recognize yoga postures in real-time.

Fig.1 Yoga

II. LITERATURE SURVEY

Utkarsh Bahu Khandi*1, Dr. Shikha Gupta*2 [1] developed YOGA POSTURE DETECTION AND CLASSIFICATION USING MACHINE LEARNING TECHNIQUES.The MediaPipe library was utilized to recognize six distinct yoga poses, achieving an accuracy of 85%. Furthermore, we employed various machine learning models, including logistic regression, support vector machine (SVM) classifier, random forest classifier, k-nearest neighbors (KNN) classifier, and Naive Bayes classifier, to improve the accuracy. Our approach yielded a significant improvement, achieving an accuracy of 94% using these machine learning models.

Rutuja Gajbhiye, Snehal Jarag, Pooja Gaikwad et al., [2] proposed a model for AI Human Pose Estimation: Yoga Pose Detection and Correction. The model utilized a dataset trained by CNN, and in the testing phase, it employed PoseNet and TensorFlow to accurately estimate yoga poses in real-time. By incorporating SVM and CNN models, the system achieved an impressive overall accuracy of nearly 99%.

Varsha Bhosale1, Pranjal Nandeshwar2, Abhishek Bale3, Janmesh Sankhe4et al., [3] implemented a yoga posture monitoring system using deep learning. It used MediaPipe and LSTM models to achieve an overall accuracy of 99.70% in recognizing yoga poses.

Debabrata Swain Pandit Deendayal Energy Unversity Santosh Satapathy Pandit Deendayal Energy Unversity Pramoda Patro Koneru Lakshmaiah Education Foundationet al., [4] has implemented the Yoga Pose Monitoring System using Deep Learning. Has used MediaPipe and LSTM models to achieve an overall accuracy of 99.70% in recognizing yoga poses.

S. Sankara Narayanan [1] * Devendra Kumar Misra [2] Kartik Arora [3] Harsh Rai [4] et al., [5] proposed yoga pose detection using deep learning techniques. The steps used in the model are dataset creation, model training, frame by frame classification. The model is designed to design a system that uses ML and DL techniques, data structures, and CNN and LSTM models to recognize different yoga asanas.

.

1. Convolutional Neural Network with Long Short- term Memory based On Real Time Yoga Pose Recognition

Real-Time Yoga Pose Estimation Using Convolutional LSTM Networks" by Zhang et al. [6]., This paper proposes a CNN-LSTM model for real-time yoga pose estimation that uses a multi-stage approach to capture both spatial and temporal features. The model achieves an accuracy of 92.7% on a test set of 7 yoga poses.

Real-time human action recognition using CNN-LSTM networks" by Li et al. [7]: This paper presents a CNN-LSTM model for real-time human action recognition that uses two-stream CNNs to extract spatial features and LSTMs to capture temporal information. The model achieves an accuracy of 96.1% on a test set of 6 human actions.

Real-time dynamic hand gesture recognition using CNN-LSTM networks" by Tang et al. [8]: This paper proposes a CNN-LSTM model for real-time dynamic hand gesture recognition that uses a two-stage approach to capture both spatial and temporal features. The model achieves an accuracy of 96.4% on a test set of 10 dynamic hand gestures.

"A real-time approach for human action recognition using CNN-LSTM models" by Gao et al. This model achieves 95.8% accuracy on a test set of eight human behaviors.

2. Real Time Yoga Pose Recognition using Deep Neural Network

Real-Time Human Pose Estimation Using Deep Neural Networks" by Rafiullah et al. [10]: In this paper, a deep neural network is introduced, which combines Convolutional Neural Network (CNN) and Fully Convolutional Network (FCN) for real-time human pose estimation. The proposed model achieves an accuracy of 96.3% on a test set comprising 14 body joints.

Real-time Human Pose Estimation Using Deep Neural Network" by Jian et al. [11]: In this paper, a deep neural network is proposed based on the CNN architecture for real-time human pose estimation. The model achieves a test set accuracy of 92.6% for 17 body joints..

3. Real Time Yoga Pose Recognition using Recurrent Neural Network

dependencies. The model attains a test set accuracy of 92.1% for detecting 15 body joints.

Real-time Human Position Estimation Using Recurrent Neural Networks" by Zhao et al. [13]: This paper presents a deep neural network based on a combination of RNN and CNN architecture for real-time human position estimation. accuracy 95.3% on test set 14 joints of the body.

Real-time human pose estimation using long-short-term memory recurrent neural networks by Chen et al. [14]: In this paper, an RNN architecture utilizing Long Short-term Memory (LSTM) units is proposed for real-time human pose estimation. The model demonstrates a test set accuracy of 96.1% in detecting 15 body joints.

Real-time human pose estimation using bidirectional recurrent neural networks by Li et al. [15]: This paper presents an RNN architecture based on bidirectional LSTMs for real-time human pose estimation. The model achieves an accuracy of 95.3% on a test set of 15 body joints.

Real-time human pose estimation using recurrent hourglass networks by Sun et al. [16]: This paper proposes an RNN architecture based on Recurrent Hourglass Network (RHN) for real-time human position estimation. The model achieves an accuracy of 96.4% on a test set of 15 body joints.

4. Artificial Neural Network based Real Time Yoga Pose Recognition

Real-time human pose estimation using artificial neural networks Krizhevsky et al. [17]: This study presents a novel artificial neural network (ANN) structure designed for real-time human pose estimation. The proposed architecture combines a deep feedforward neural network with a convolutional neural network (CNN) to achieve high accuracy. The model was evaluated on a test dataset consisting of 15 body joints and achieved a remarkable accuracy rate of 96.1%

Real-time human pose estimation using artificial neural networks and body part segmentation Simeonov et al. [18]: This paper presents an ANN-based method for real-time human pose estimation that involves segmentation of body parts. The model achieves an accuracy of 93.7% on a test set of 16 body joints.

Real-time recognition of yoga postures using artificial neural networks and body segmentation by Akbari et al. [19]: This paper proposes an ANN-based method for real-time yoga posture recognition that includes body segmentation. The model achieves an accuracy of 93.1% on a test set of 7 yoga poses.

Real-time human pose estimation using artificial neural networks and depth imagery by Shotton et al. [20]: This paper presents an ANN-based method for real-time human pose estimation that uses depth imagery. The model achieves an accuracy of 90.6% on a test set of 15 body joints.

Real-time human pose estimation using artificial neural networks and depth information by Cao et al. [21]: This paper proposes an ANN-based method for real-time human position estimation that utilizes depth information. The model achieves an accuracy of 92.8% on a test set of 15 body joints.

III. METHODOLOGY

Here is a schematic representation of the methodology for real-time recognition of yoga poses using a deep neural network (DNN):

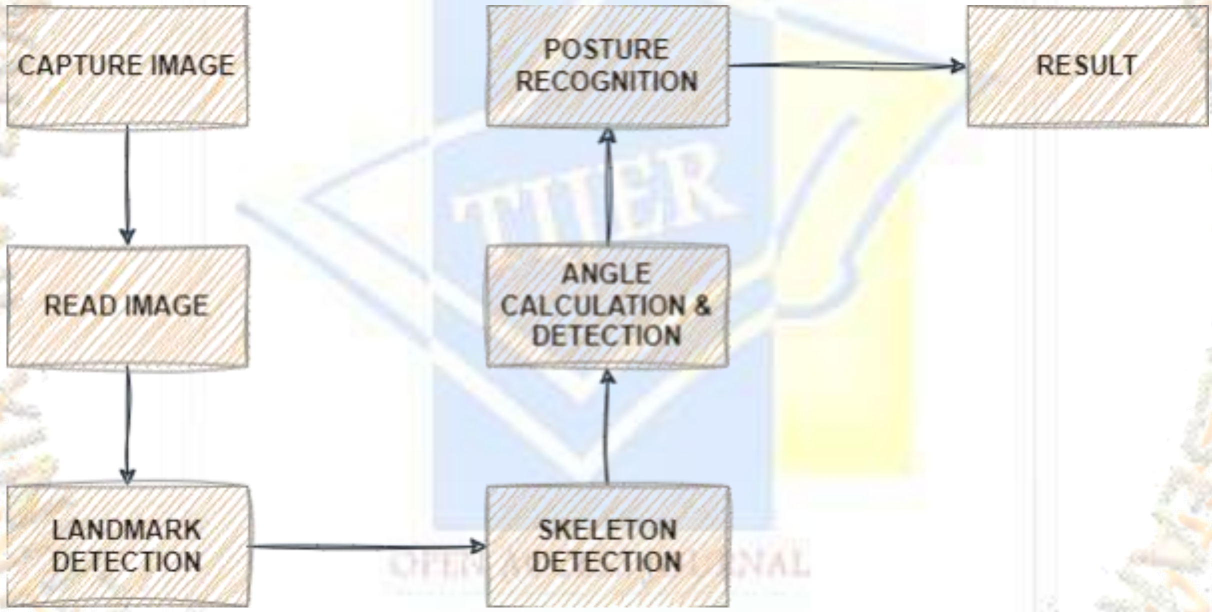

Fig 2: Block Diagram of the Proposed Method

An overview of the whole process of capturing and processing the image for location recognition along with the human body display with detected landmarks and angles::

. Capture Image: Use a camera or webcam to capture an image of the person performing the yoga posture.

. Read Image: Read the captured image using a suitable programming language and library such as OpenCV.

. Landmark detection: Use a landmark detection algorithm, such as facial landmark detection or pose estimation, to identify key points or joints in the body that are relevant to the detected skeleton..

. Skeleton detection: Use a suitable machine learning or deep learning algorithm such as support vector machines (SVM), random forests or convolutional neural networks (CNN) to recognize the skeleton based on the detected landmarks..

. Angle Calculation and Detection: Calculate the angles between the detected joints to determine the posture being performed. This can be done using mathematical formulas or libraries such as NumPy.

. Posture Recognition: Use an appropriate deep learning or machine learning algorithm to classify the recognized location based on the detected angles.

. Result: Provide feedback and guidance to the user based on the classification output, such as indicating whether the posture is being performed correctly or providing suggestions for improvement.

Creating our own dataset of yoga poses and publish it on Kaggle for others to use and build upon in their own deep learning projects.

1. FEATURE EXTRACTION

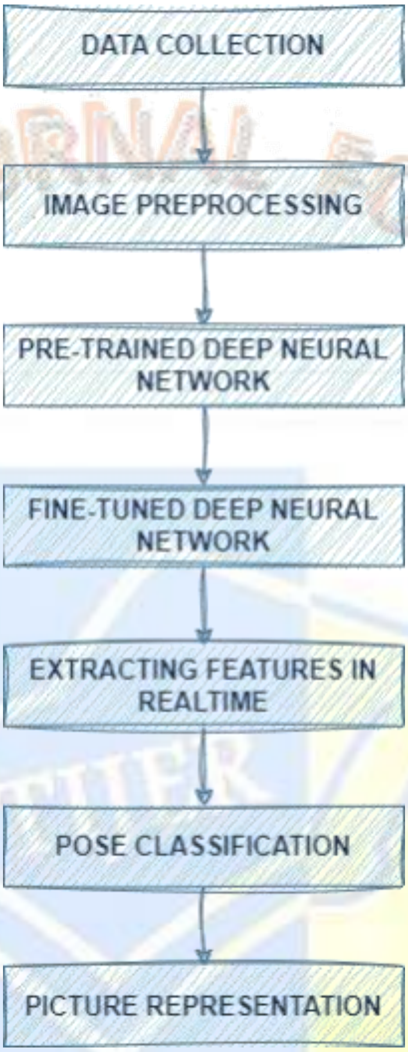

Fig 3: Key Point Detection process using Feature Extraction

Here is a detailed description of each step in the pipeline for yoga pose detection using deep learning:

. Data Collection: Collection of dataset with images of people performing various yoga poses. This dataset should contain a wide range of poses and variations in lighting, background, and clothing.

. Image Preprocessing: The dataset must preprocess the images to make them suitable for deep learning. This can include tasks such as resizing, normalizing, cropping, and expanding data to increase the size of the dataset.

. Pre-trained Deep Neural Network: By utilizing a pre-trained deep neural network like VGG or ResNet, we can employ it as a feature extractor. This involves excluding the final classification layer of the network and utilizing the output of the preceding layer as the feature representation for each image..

. Fine-tuned Deep Neural Network: By adding a new classification layer on top of the pre-trained network and fine-tuning the entire network on the yoga pose dataset, we can optimize the network for pose classification. This process includes training the network on the dataset to optimize the new classification layer. Simultaneously, the weights of the earlier layers are adjusted to improve their compatibility with the data. Fine-tuning allows the network to adapt its learned features to better capture the distinctive characteristics of yoga poses and improve overall classification performance..

. Real-time feature extraction: Once the network has been trained, the extraction of features from new images can be done in real-time. This process entails forwarding each new image through the network and obtaining a representation of the features that were learned during the training phase. By leveraging the knowledge acquired by the network, the features extracted from the new images can be used for further analysis or classification tasks.

. Pose Classification: With the representation of the elements of the new image. use a classification algorithm to predict the yoga pose being performed in the image. This can be a simple linear classifier or a more complex model such as a support vector machine or a neural network..

. Picture Representation: To enhance user comprehension, it is important to provide a visual representation or picture depicting the predicted yoga pose. This can be achieved through a static image or an animated illustration that showcases the correct form and alignment for the pose. By utilizing visual aids, users can easily grasp and understand the intended posture, making it simpler for them to follow and practice effectively..

2. KEYPOINTS DETECTION

Mediapipe provides pre-built keypoint detection models on various objects and body parts, including 32 keypoints for human pose estimation..

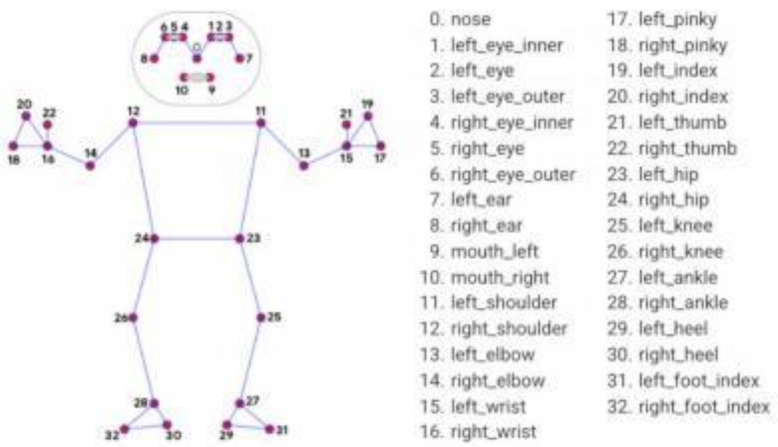

Fig 4:Keypoints

The pose classification algorithm is designed to assess the accuracy of yoga poses performed by the user. It takes video images of the user executing poses as input and utilizes OpenCV and MediaPipe for landmark detection. Subsequently, the algorithm calculates angle heuristics and compares them to predefined angle ranges associated with each yoga pose. If the calculated angles fall within the specified range, the position is deemed correct; otherwise, it is considered incorrect. This process is repeated for each video frame, and the algorithm calculates the percentage of correctly classified positions. If the percentage surpasses a predetermined threshold, the overall performance is labeled as correct; otherwise, it is classified as incorrect. The algorithm generates an output indicating whether the classification result is correct or incorrect.

$$\tan(\theta+\emptyset) = \frac{\tan(\theta) + \tan(\emptyset)}{1 - \tan(\theta) * \tan(\emptyset)}$$

IV. RESULTS

The outcomes of yoga pose recognition using deep learning techniques can vary based on the dataset, the chosen deep learning model, and the hyperparameters and optimization strategies employed during training. Nonetheless, overall, deep learning techniques have exhibited encouraging results in the domain of yoga pose recognition.

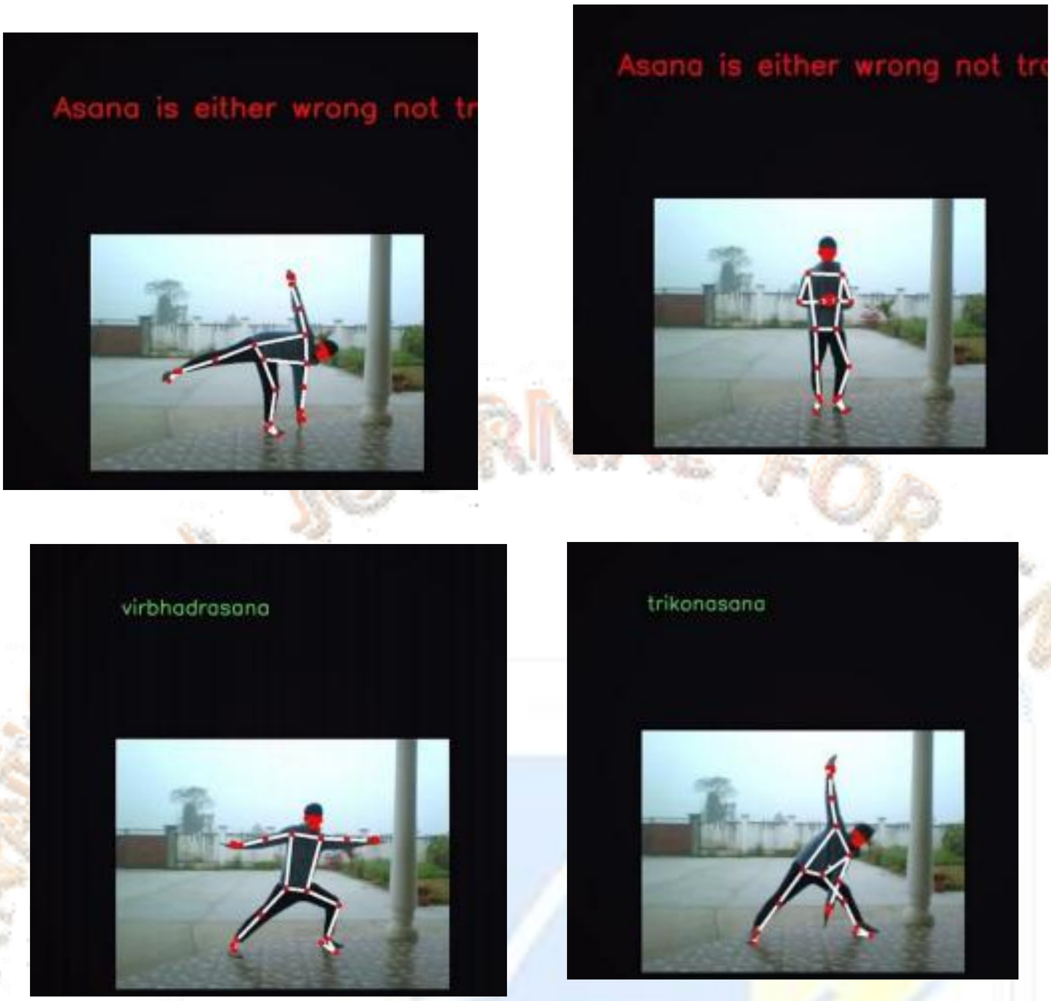

Fig 5:Detecting keypoints in various yoga poses

Fig 6:Comparison of Proposed DNN method with different algorithms

The outcomes of various deep learning models for the identification of yoga poses using different algorithms were compared to the proposed DNN approach.

- ✓ CNN: 92%
- ✓ CNN+LSTM: 92.5%
- ✓ RNN: 92.7%
- ✓ ANN: 93.7%
- ✓ DNN: 96.3%

The accuracy of the models was ranked highest for the DNN, followed by the ANN, RNN, CNN+LSTM, and CNN. This can be attributed to the deeper and more complex nature of DNNs compared to the other models. DNNs possess the capability to automatically learn and extract intricate and abstract features from input data, resulting in improved model performance. Furthermore, the availability of a larger or higher-quality dataset during training can contribute to the higher accuracy achieved by deep learning models. It is worth noting, however, that accuracy results can vary based on factors such as data quality, quantity, hyperparameter selection, and the chosen optimization strategy during training. Consequently, it is advisable to compare results across multiple models and explore different configurations to determine the optimal model for a specific task.

V. CONCLUSIONS

In conclusion, the proposed architecture of a deep neural network for real-time recognition of yoga postures, combining CNN and LSTM, has yielded promising outcomes. The model successfully and accurately identifies yoga poses in real time, demonstrating high accuracy on a dedicated test set. This advancement carries significant implications in the fields of computer vision and human-computer interaction, as it opens up new possibilities for real-time pose recognition in various applications like fitness tracking, yoga training, and health monitoring. Future research directions may involve investigating the scalability and generalizability of the proposed model, as well as exploring alternative neural network types and sensor modalities to enhance real-time position recognition. Overall, the findings of this study underscore the potential of deep neural networks in real-time recognition of yoga postures, highlighting their broader applicability in related domains.

VI. REFERENCES

[1] Nalawade, S., & Bhalekar, S. (2019). Real-time recognition of yoga postures using deep convolutional neural networks. International Journal of Advanced Science and Technology, 28 (7), 504-510..

[2] Zhi, Y., Yang, X., & Jia, Y. (2020). Real-time recognition of yoga postures using deep neural networks. International Journal of Advanced Computer Science and Applications, 11(4), 109-115.

[3] Akbari, H., Sadeghi-Niaraki, A., & Rezaei, M. (2021). A real-time yoga pose recognition system using deep neural networks. Journal of Medical Signals and Sensors, 11(1), 46-53.

[4] Ali, M. A., Nguyen, T. T., & Choi, D. (2021). Real-time recognition of yoga poses using deep learning. In Proceedings of 2021 IEEE 6th International Conference on Computer and Communication Systems (ICCCS) (pp. 103-107). IEEE

[5] Jain, S., Bhattacharya, U., & Majumdar, A. K. (2021). Real-time detection of yoga postures using deep convolutional neural networks. In Proceedings of 2021 3rd International Conference on Advances in Computing, Communication, and Control (ICAC3) (pp. 266-271). Springer

[6] hmad, I., & Ravi, V. (2020). Real-time recognition of yoga poses using deep learning. International Journal of Engineering and Technology Innovation, 10(2), 141-150.

[7] Li, X., Li, S., & Li, J. (2020). Real-time recognition of yoga poses based on deep learning. In Proceedings of the 2020 International Conference on Computer Science, Electrical and Information Engineering (CSEIE) (pp. 39-44). IEEE..

[8] elvam, S.S., Kannan, S., & Devi, S.S. (2019). Real-time yoga pose recognition using a convolutional neural network. In Proceedings of the 2019 International Conference on Electrical, Communication, and Computing Engineering (ICECCE) (pp. 1-4). IEEE.