Minor Project Report

on

“Explorations in Yoga Pose Detection using CV Models”

by

Kushagra Saruparia (20BEC027), Rushay Modi (20BEC032), Vaibhavi Udgirkar(20BEC038), Harsh Shah (20BEC045), Dev Pandya (20BIT138)

Under the Guidance of:

Dr. Ritesh Vyas

Assistant Professor, Department of ICT

Submitted to:

School of Technology,

Pandit Deendayal Energy University

CERTIFICATE

This is to certify that the minor project report entitled "Explorations in Yoga Pose Detection using CV Models," submitted by Kushagra Saruparia, Rushay Modi, Vaibhavi Udgirkar, Harsh Shah & Dev Pandya, has been conducted under the supervision of Dr. Ritesh Vyas and is hereby approved for the partial fulfillment of the requirements for the award of the degree of Bachelor of Engineering in the Department of ECE & Department of ICT at Pandit Deendayal Energy University, Gandhinagar. This work is original and has not been submitted to any other institution for the award of any degree.

Sign:
Name of the Guide: Dr. Ritesh Vyas
Designation: Assistant Professor
Department: Information & Communication Technology (ICT)
School of Technology
Pandit Deendayal Energy University

Sign:
Name of Examiner:
Designation:
Department:
School of Technology
Pandit Deendayal Energy University

DECLARATION

We hereby declare that the seminar report entitled "Explorations in Yoga Pose Detection using CV Models" is the result of our work and has been written by us. This report has not utilized any language model or natural language processing artificial intelligence tools for the creation or generation of content, including the literature survey.

The use of any such artificial intelligence-based tools was strictly confined to the polishing of content, spell-checking, and grammar correction after the initial draft of the report was completed. No part of this report has been directly sourced from the output of such tools for the final submission.

This declaration is to affirm that the work presented in this report is genuinely conducted by us and to the best of our knowledge, it is original.

Kushagra Saruparia
20BEC027
Department of ECE

Rushay Modi
20BEC032
Department of ECE

Vaibhavi Udgirkar
20BEC038
Department of ECE

Harsh Shah
20BEC045
Department of ECE

Dev Pandya
20BIT138
Department of ICT

School of Technology
Pandit Deendayal Energy University
Gandhinagar

Date:
Place: Gandhinagar

List of Tools Used for the Report with Purpose:
- Grammarly: Correcting Grammar.
- ChatGPT: Polishing the text.

ACKNOWLEDGEMENT

We wish to extend our sincerest appreciation to everyone who has played a vital role in the triumphant culmination of our Minor Project “Explorations in Yoga Pose Estimation using CV models”. First and foremost, we would thank Dr. Ritesh Vyas our mentor, for his invaluable guidance, unwavering support, and constructive feedback throughout the research and development process. His expertise and encouragement have been instrumental in shaping our understanding of the subject matter and enhancing our skills.

Additionally, we extend our heartfelt gratitude to PDEU (Pandit Deendayal Energy University), our university that provided us with the platform, resources to conduct this research, collaborative and welcoming environment that fostered our growth.

Furthermore, we extend our sincere appreciation to all the individuals who have contributed to our growth and learning during this project. We are truly grateful for their guidance and support. This endeavor has proven to be a momentous expedition of acquiring knowledge and skills in the domain of Computer Vision models and Web Development.

Lastly, we acknowledge and appreciate the collaborative effort, dedication, and unique insights each team member brought to the project, making it a collective success.

Sincerely,
Kushagra Saruparia (20BEC027)
Rushay Modi (20BEC032)
Vaibhavi Udgirkar (20BEC038)
Harsh Shah (20BEC045)
Dev Pandya (20BIT138)

ABSTRACT

Yoga is an ancient practice that originated in ancient India during the time of the Indus-Sarasvati civilization, which dates back 5000 years. The term yoga signifies a profound union and connection between the mind and body. Its purpose is to achieve an union state of body mind and soul through the practice of asana, meditation, and various other techniques. In this progressive era, individuals are relentlessly pursuing and aspiring for success, and in order to achieve this, they work diligently.

Exercises including yoga poses are required to be performed properly without any deformity in posture which leads to injuries. If there is any anomaly, it can cause a hazardous and life-threatening impact on a person's life. To mitigate this problem, we require an accurate, precise, pose detection model that can help in guiding and directing an individual towards the correct practice of performing Yoga. Leveraging the facility our model provides one can learn the skill at their own pace without the intervention of any trainer which also feels like a treat to their inhibitions.

The project aims to present an overview of the progress and application of a groundbreaking Yoga Pose Detection system that leverages MediaPipe technology to analyze real-time live video streams. The main goal is to establish a dynamic and user-friendly platform that enables people to attain accurate yoga poses in real-time. The system goes beyond just technical precision, incorporating imaginative and captivating elements to enrich the overall yoga journey.

"Exploration in Yoga Pose Detection using CV Models" project seeks to develop a harmonious fusion of technology and mindfulness, revolutionizing how people interact with and benefit from their yoga practice by fusing technical precision with creativity.

TABLE OF CONTENTS

Sr. No.	Content	Page No
1.	Abstract	5
2.	List of Tables	7
3.	List of Figures	8
4.	Chapter1: Introduction	10
5.	Chapter 2: Literature Survey	12
6.	Chapter 3: Methodology	17
7.	Chapter 4: Procedure	23
8.	Chapter 5: Result Analysis & Discussion	27
9.	Conclusion	33
10.	Future Scope	34
11.	Refrences	35
12.	Plagiarism Report	36

LIST OF TABLES

Sr.No	Table	Page No
1.	Comparison between Mediapipe and Movenet	25
2.	Asanas & Model Accuracies	28

LIST OF FIGURES

Figure Number	Figure
Figure 1.	Flowchart of Methodology
Figure 2.	Architecture of MediaPipe
Figure 3.	Correct Posture of Plank to train the model
Figure 4.	Output with skeleton view and key points for plank.
Figure 5.	Correct Posture of Warrior asana to train the model
Figure 6.	Output with skeleton view and key points for warrior asana
Figure 7.	Flowchart of Procedure
Figure 8.	Correct posture for Tree pose mapped in green color
Figure 9.	Incorrect posture for Tree pose mapped in red color
Figure 10.	Mapped contours of Downdog Pose
Figure 11.	Mapped contours of Goddess Pose
Figure 12.	Mapped contours of Tree Pose
Figure 13.	Mapped contours of Plank Pose
Figure 14.	Mapped contours of Warrior Pose
Figure 15.	User interface of website
Figure 16.	About section of the website.
Figure 17.	Interface after the user selects an asana Vrikshasana
Figure 18.	Interface after the user selects an asana Phalakasana
Figure 19.	Live feed for incorrect posture for tree pose being performed

Figure 20.	Live feed for correct posture for tree pose being performed
Figure 21.	Live feed for incorrect posture for Goddess pose being performed
Figure 22.	Live feed for correct posture for Goddess pose being performed

Chapter 1: INTRODUCTION

In a period characterized by inactive ways of life and progressions in technology, the significance of upholding one's physical welfare is now more imperative than it has ever been. There is no denial in the fact that physical and mental fitness is an essential aspect of acing any work as it upsurges the concentration, coordination between the mind and body, reliability, self-esteem, etc.

Yoga, being an all-in-one exercise, is highly rated by the people. But as said, to do something effectively, one must do it in the right way. Therefore, it becomes quintessential to do yoga properly and cautiously to avoid injuries and abnormalities. To ensure the correct way of doing yoga, people often require a guide or a trainer since yoga is not something congenital. Not everyone has resources like a gym and cannot afford trainers. AI system having a ability to classify different posture in real time with benefit of providing feedback to the postures would be of great help to performer.

In this project, we have focused on estimating and comparing different models, thus, inferring the model that suits best for a certain yoga pose estimation. This will aid in the selection of a robust and efficient model for the later stage of the project. A system provide feedback and key improvement to make sure to have proper and injury free yoga postures.

The individuals have the capability to choose the preferred position for exercise and have the capability to execute it in real-time in the presence of a webcam. The position of the individual is transmitted to train models that yield the anomalous angles identified between the actual position and the individual's position. Utilizing these yields, the system provides recommendations to the individual on how to enhance the position by specifying the areas in which the yoga position is being executed incorrectly. Following the acquisition of an inaccurate position, the individual has the ability to identify and correct their position, thus enhancing their posture. We aim to effectively compare the subsequent models: MediaPipe[1] and MoveNet[9].

Key Objectives:

- **Real-time Pose Detection:** Design and implement a robust system that is capable of detecting and tracking yoga poses in real-time using the MediaPipe framework.
- **Feedback and Correction:** Provide users with immediate feedback on their pose, which will allow them to make necessary adjustments in order to improve their posture.
- **Pose Recognition and Labelling:** Employ algorithms to recognize and label common yoga poses, thereby assisting users in comprehending and attaining the correct form.
- **User-Friendly Interface:** Develop an intuitive interface that facilitates seamless interaction, ensuring accessibility for users of all proficiency levels.
- **Alignment Guidance:** Provide precise guidance on alignment to minimize the risk of injuries and optimize the effectiveness of yoga practice.

Chapter 2: LITERATURE SURVEY

1. *Real-time Yoga recognition using Deep learning. [6]*

The research article introduces a real-time system for recognizing yoga movements using deep learning algorithms. Specifically, a hybrid model consisting of a CNN and a long short-term memory (LSTM) is employed. The CNN layer is responsible for extracting features from key points obtained through the utilization of OpenPose. On the other hand, the LSTM layer is responsible for making predictions based on the temporal information. By employing this proposed system, the necessity for manually crafted features is eliminated, and the addition of new yoga poses is made possible by retraining the model with new data. The accuracy of the system is reported to be 99.04% on individual frames and 99.38% after aggregating predictions from 45 frames of the videos. Additionally, the system is tested in real-time with a different group of 12 individuals, achieving an accuracy of 98.92%. The dataset used in this study consists of 15 individuals, each performing six selected yoga poses.

The methodology of this study encompasses four major steps: data collection, pose extraction using OpenPose, feature extraction using CNN, and temporal analysis using LSTM. OpenPose, an open-source library, is employed for pose extraction, enabling the detection of key points related to the human body, hands, and face. The proposed system utilizes OpenPose for extracting key points, followed by the employment of CNN and LSTM models for predicting the user's yoga poses.

In the hybrid model, the CNN layer is responsible for extracting spatial features from the key points obtained through OpenPose. On the other hand, the LSTM layer analyzes the temporal changes in these features across frames, capturing the sequential nature of yoga movements. The model employs thresholding to identify frames where the user is not performing yoga. The implementation of the model is carried out using Keras Sequential API in Python and achieves a test accuracy of 99.04% on individual frames and 99.38% after aggregating predictions from 45 frames. By incorporating LSTM into the model, information from previous frames is leveraged, leading to more accurate and robust predictions. Moreover, the system is tested in real-time with a different set of 12 individuals, resulting in an accuracy of 98.92%.

2. *A Proposal of Yoga Pose Assessment Method Using Pose Detection for Self-Learning [2]*

The present study proposes a methodology for evaluating Yoga postures through the use of pose detection in order to facilitate self-directed learning. The efficacy of this methodology was assessed by observing three individuals executing fundamental Yoga poses, and it successfully identified incorrect aspects of each pose. The detection of poses was accomplished through the utilization of OpenPose in conjunction with a PC camera, whereby the discrepancy in body angles between the instructor and the user was computed, and corrections were suggested if necessary. The objective of this approach is to empower individuals to engage in Yoga practice in any setting, including the home environment, irrespective of age or health status. OpenPose, which is employed for Human Pose estimation, serves to localize anatomical key points or components, with a primary emphasis on identifying the various body parts of the individual. This is achieved through the utilization of part affinity fields and a multi-convolutional neural network (CNN) architecture, which facilitate the extraction of human body parts and their associations with a high degree of accuracy.

The development of OpenPose was predicated on the joint acquisition of parts detection and parts association. The connection process is guided by the part affinity score. The central premise of this proposal centers on the measurement of similarity between the angles of specific body parts of the learner and those of the instructor. The authors of this study computed the relevant parameters and angles between the body joints, which provided a precise depiction of the location and contours within the input image or video. To enhance the comprehensibility of the outcomes, the developers have incorporated a gradient color display. Each joint is assigned an angle difference value, which corresponds to a particular color value. Specifically, red is utilized to signify a large angle difference, indicative of an erroneous outcome, while green is employed to denote a small difference, indicative of a correct outcome. In order to promote self-directed learning of Yoga, this investigation has devised a system for evaluating performance, which serves as a training tool for Yoga postures.

The system assesses an individual's Yoga pose by identifying the pose or skeleton, determining the discrepancy in body angles between the instructor and the user, pinpointing the incorrect portion between the two, and categorizing the pose into four levels based on the average angle difference. The effectiveness of this methodology was demonstrated through its application to three individuals of varying ages, genders, and body types, who performed three different Yoga positions. Consequently, the system's efficacy was substantiated. Furthermore, the system has the potential to be applied to various physical exercises, thereby enabling the evaluation of exercise poses in future research endeavors.

3. *Yoga pose classification: a CNN and MediaPipe inspired deep learning approach for real-world application [5]*

The paper examines the research conducted in the domain of yoga pose classification utilizing deep learning methodologies and accentuates the challenges encountered in the classification of yoga postures, as well as the significance of precise posture detection in yoga practice. Numerous investigations have been carried out to devise systems for yoga pose classification. For instance, Chen et al. (2014) proposed YogaST, a posture correction system that employed kinect technology to categorize poses. Patil et al. (2011) introduced the yoga tutor, which identified disparities in posture formation. Sukarsa et al. (2020) employed neural networks to identify crucial points in the human body for yoga pose classification.

Other researchers have explored the utilization of deep learning models for yoga pose classification. Byeon et al. (2020) presented an amalgamated deep model based on multiple Convolutional Neural Network (CNN) models. Agrawal et al. (2020) introduced the YOGI dataset and employed Deep Learning Convolutional Neural Network (DLCNN) for posture recognition. Cao et al. (2021) proposed a technique to detect poses in videos featuring multiple individuals. The review also highlights the challenges in yoga pose estimation, such as the insufficiency of adequate datasets and the necessity for essential point detection techniques.

It emphasizes the employment of MediaPipe, a library developed by Google, for real-time detection of yoga postures. MediaPipe employs a detector and a tracker to identify skeletal keypoints in humans. Concerning deep learning models, the review examines the use of the Inception-v3 and VGG16 architectures. The Inception-v3 architecture is a transfer-learning architecture consisting of 159 layers, whereas VGG16 is a more intricate architecture with a size of 533 MB. These models have been utilized for yoga pose classification, yielding favorable outcomes. The document also presents the proposed model, YogaConvo2d, which is constructed utilizing 2D convolutional neural networks. The model achieved an accuracy of over 90% in the classification of five distinct yoga pose classes. Overall, the literature review accentuates the significance of precise yoga pose classification and the progress made in this field through the utilization of deep learning methodologies.

4. *Yoga Pose Estimation and Feedback Generation Using Deep Learning [7]*

Yoga is an ancient practice that originated in India around 5,000 years ago during the Indus-Sarasvati civilization. It aims to achieve equilibrium of the body and mind through the practice of asanas, meditation, and various other techniques. Poor posture can have adverse effects on health, causing both acute pain and long-term chronic issues. This article presents the development of deep learning-based techniques for identifying incorrect yoga postures. By utilizing this approach, individuals can select their desired pose and upload recorded videos of their yoga practices. The user's posture is then analyzed by training models that generate abnormal angle measurements between the actual posture and the user's posture.

The proposed method has been compared to several advanced techniques and has demonstrated exceptional accuracy, achieving a score of 0.9958, while also requiring less computational complexity. The researcher conducted a comprehensive review of relevant literature on the detection and correction of yoga poses, identifying their limitations and challenges. This led to the development of a more precise model, aimed at minimizing any disadvantages based on informed judgement. When detecting human posture, various factors are taken into consideration, including the environment, human interaction, and clothing variations.

The system employs classifier networks, such as multilayer perceptron, along with hyperparameter tuning, to achieve high accuracy. The method consists of three main stages: feature extraction, classification, and response generation. The dataset utilized is open source and comprises 70 videos with a total of 350 instances of 6 yoga poses, namely Cobra (Bhuj), Tree (Vriksh), Mountain (Tada), Lotus (Padam), Tam Sense (Trik), and Corpse (Shav). Once the image is inputted into the model, the trained model classifies the pose it corresponds to. The extracted angles from the image are then compared to the calculated mean values.

To provide recommendations, two parameters are required: the degree to which the posture deviates from the original and the direction of deviation. The magnitude of these 12 offsets indicates the extent of postural correction needed, while the positive or negative sign indicates the direction in which the joints should be rotated (clockwise or counter-clockwise). Using this method, suggestions are offered to the user for each match. The proposed method maintains a low computational complexity, making it suitable for busy individuals who want to incorporate personal yoga learning into their lives. Furthermore, it effectively detects incorrect yoga postures, helping to prevent chronic issues.

5. *A novel approach for yoga pose estimation based on in-depth analysis of human body joint detection accuracy [4]*

Human pose estimation poses a significant challenge in the field of computer vision, particularly when applied to the recognition and classification of yoga poses. Researchers have explored various machine learning frameworks to detect body joints, including models based on wearable sensors and approaches utilizing the Kinect technology. However, these methods have limitations in terms of their convenience and practicality. Recent research has introduced the BlazePose model as a highly effective solution for accurately detecting body joints. This model employs a mesh model to represent the 3D pose estimation and a skeleton model for localizing body joints.

The accuracy of estimated body joints is evaluated using parameters such as the percentage of detected joints (PDJ) and the percentage of corrected keypoints (PCK). The precise detection of body joints is essential for accurate pose estimation and classification, particularly in the context of yoga exercises where the localization of body joints holds significant importance. Existing techniques for identifying yoga poses have primarily focused on deep neural network-based classification models, evaluating their performance using parameters derived from the confusion matrix. However, an alternative approach involves assessing yoga pose classification based on the accuracy of joint detection, as measured by parameters like PCK and PDJ. While previous studies have utilized the OpenPose model for yoga pose classification, this model does not evaluate the accuracy of body joint localization according to standard evaluation parameters.

To overcome these limitations, researchers have adopted the BlazePose model for accurate body joint detection in yoga poses. This model has the ability to detect 33 human body joint landmarks, including those in the head, torso, arms, and legs, in real-time. The accuracy of pose estimation is evaluated using parameters such as PCK and PDJ. The PDJ parameter takes into account the mean distance between predicted and ground truth joints, while the PCK parameter assesses the distance within specific threshold limits. The results demonstrate that the model achieves high accuracy in detecting body joints, although some poses exhibit better accuracy than others. In comparison to other models such as OpenPose and HrNet, the BlazePose model outperforms them by detecting a greater number of body joints and achieving superior accuracy in pose estimation. Thus, the BlazePose model has emerged as a promising solution, displaying high accuracy in both body joint detection and yoga pose estimation. The evaluation parameters, PCK and PDJ, provide quantitative measurements of accuracy and can be utilized to evaluate the performance of different pose estimation models.

.

Chapter 3: METHODOLOGY

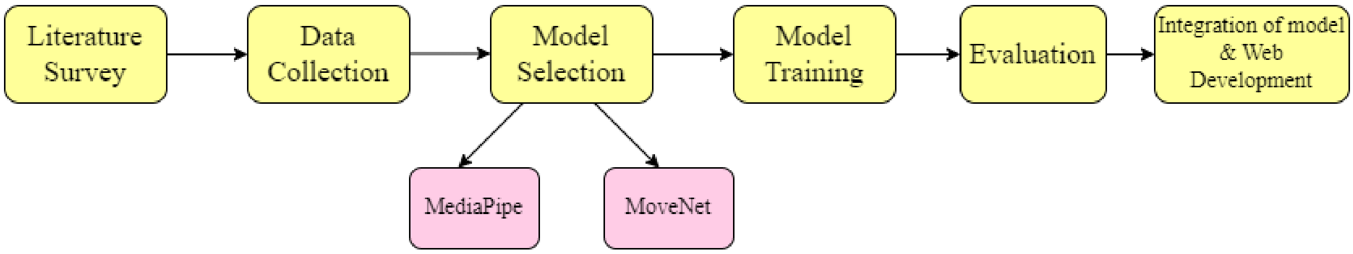

Figure 1: Flowchart of Methodology

1. Data Collection:

Capturing the subtleties of each pose, such as body alignment, breath control, and mindfulness techniques, is largely dependent on experienced yoga practitioners and instructors.To enable accurate model learning, comprehensive data regarding joint angles, body orientation, and important anatomical points are carefully captured. Six thousand photos representing the five different asanas—the downdog, goddess, plank, tree pose, and warrior pose—were included in the dataset[3] [8]. This aids in both the model's training and the quantized data calculation.

2. Model Selection:

The model selection part entailed researching and reviewing various models like OpenPose, MediaPipe, MoveNet, PoseNet, etc. from which we selected two models for our project. A holistic conducted an extensive literature review on MoveNet and MediaPipe, studying their architectures, performance, and applications in pose estimation tasks.

- **MediaPipe[1]**

  MediaPipe is an open-source framework developed by Google that provides tools and pre-trained models for various perceptual computing tasks, including pose estimation. The framework allows developers to build cross-platform applications with customizable pipelines for real-time processing. MediaPipe is used in audio, image, and video processing. It has features for identifying and mapping the contours of the body like the joints, face and hand landmarks, etc. It works on video frame by frame, converting it to an image which is further transformed to Tensor. The tensors are then fed into the

MediaPipe model which is then processed to detect the landmarks followed by rendering and converting it again in the video form to display it to the user.

Specifically for the estimation of yoga pose estimation, MediaPipe employs computer vision techniques and machine learning models to identify and track critical points on an individual's body as they execute various yoga poses. These critical points are subsequently utilized to determine the pose and offer feedback or analysis based on the predefined angles of correct poses, thereby facilitating the practice of yoga or the correction of form.

Through the utilization of MediaPipe's tools and models, we have developed a model that aids yoga practitioners in enhancing their poses and delivering real-time feedback during yoga sessions.

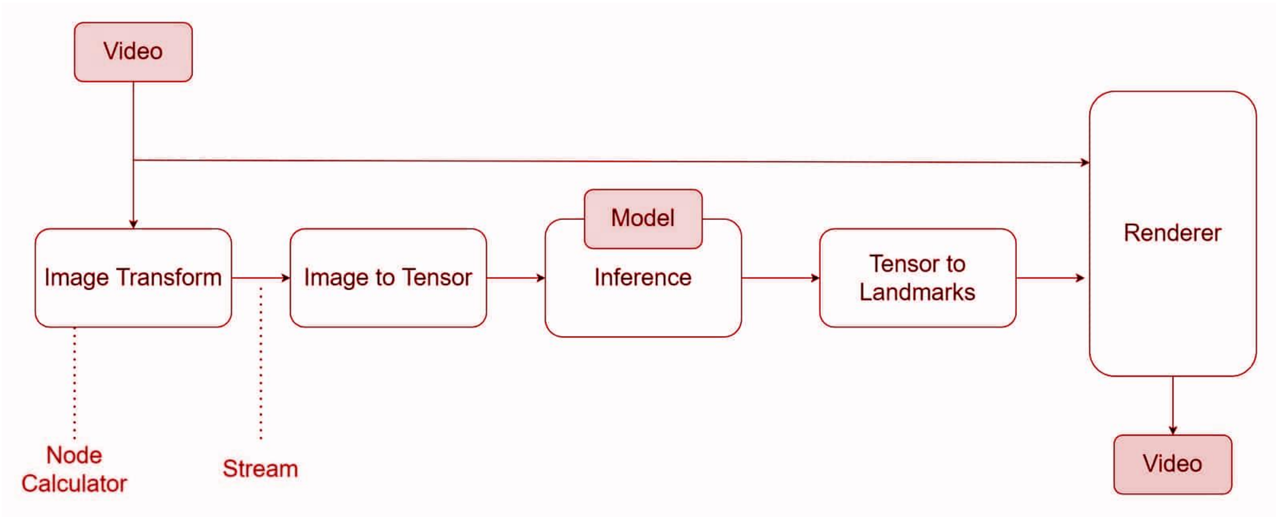

Figure 2: Architecture of Mediapipe

- **MoveNet:**

MoveNet is a lightweight machine learning model developed by Google that focuses on human pose estimation. MoveNet is specifically designed to accurately identify and monitor multiple crucial points on an individual's body in real-time using video or image inputs. This functionality allows for efficient operation across various devices, including smartphones, thereby rendering it well-suited for applications necessitating real-time pose estimation. The architecture of MoveNet has been optimized to prioritize both speed and precision when detecting human poses. Its methodology involves employing a single-shot detector approach coupled with keypoint regression in order to identify significant points such as joints and landmarks on the human body. These identified points are subsequently utilized to estimate and track an individual's pose. The model encompasses distinct versions tailored to address specific use cases, encompassing full-body pose estimation as well as hand pose estimation. The adaptability and efficiency of MoveNet render it highly suitable for a diverse array of applications, ranging from fitness tracking to gesture recognition, augmented reality experiences, and beyond, all of which necessitate real-time pose estimation.

When applied to yoga pose estimation, MoveNet meticulously scrutinizes video streams or images of the human body to discern and track distinct key points, such as joints and limbs, as an individual executes yoga postures. This functionality facilitates the monitoring and analysis of yoga postures accuracy in real-time. The model's remarkable capacity to accurately determine poses can be effectively employed in applications intended to support yoga practitioners in refining their technique and postures and offering constructive input on their poses. The model is particularly efficient and suitable for scenarios where resource constraints and low-latency are critical, making it well-suited for yoga pose estimation.

3. **Model Training:**

To identify the essential characteristics of the human body in different yoga poses, pre-trained models were utilized to recognize important body parts. The required angles for each pose were then determined using the coordinates of these keypoints, accounting for the various ranges of motion related to various yoga positions. The study methodically measured and tested these angles to determine the appropriate parameters for ideal postures using an angle calculator and keypoint mapping.

After these angles were determined, test images were classified into correct and incorrect yoga postures using datasets by a trained model. The model offered useful feedback by comparing the measured angles with established norms, which helped practitioners improve their form and guarantee that proper pose execution was followed. This comprehensive method, which combined keypoint mapping, angle computations, and pre-trained models, created a complex system that could precisely assess and classify yoga poses and added to our understanding of the proper body alignment during practice.

After experimenting with various learning rates, we found that a learning rate led to optimal performance.To determine the range of appropriate yoga poses, the model underwent 50 training epochs, with early stopping incorporated to oversee validation performance. After ten consecutive epochs of 0.001, early stopping was initiated if there was no improvement or error in the information extraction.We used data augmentation techniques, such as converting the color index from rbg to bgr format, to increase the robustness of the model. A dataset that was divided into training and testing sets with 80% and 20% splits, respectively, was used to train the model.

Figure 3: Correct Posture of Plank to train the model

Figure 4: Output with skeleton view and key points for plank.

Figure 5: Correct Posture of Warrior asana to train the model

Figure 6: Output with skeleton view and key points for Warrior asana.

4. **Evaluation:**

The following formula is used to determine the accuracy of the model using the two distinct classes of incorrect and correct postures:

$$Accuracy = \frac{No.of\ Predicted\ Correct\ Postures}{Total\ Number\ of\ Correct\ Postures} * 100$$

5. Website Development:

For introducing our built model to people, we thought it's best to put it on the internet so that anyone can access it and gain advantage from it. To realize this idea, we developed a website where we have integrated the model.

- *Front-End Design:* The website's user interface was designed with a focus on intuitiveness and user-friendliness.
- *Integration:* Trained models were seamlessly integrated into the website to enable real-time yoga pose estimation.

The website contains both the frontend and the backend. The frontend is basically the visible website and its different elements to which the users can interact. The backend is how the video frames are sent over the internet and processed for the estimation of the model. The video frames are sent every second to the server which is processed and made compatible with the model.

Django, a python framework for backend development, is used for the website's backend development. It is an open source, high-level framework for neat, clean, and pragmatic writing of the backend of websites. Django is ridiculously fast as it provides the developer with the ability to quickly transform the idea to completion. It has inbuilt security measures that makes it easier for the developer to maintain and offer security to the end users. We chose to use django as our backend framework as it provides ultimate compatibility with running machine learning models on the web and every python library can be efficiently used in django.

For the frontend design, we used HTML, CSS, and JavaScript. We have tried our best to give a calm look to the website and provide a smooth user interface.

- Feedback Mechanism: Through the website, the user can send his video to the server which is processed and sent back to the user his comprehension. The website displays the live feed onto which the user can view his mappings and correctness of the performed yoga pose.

Chapter 4: PROCEDURE

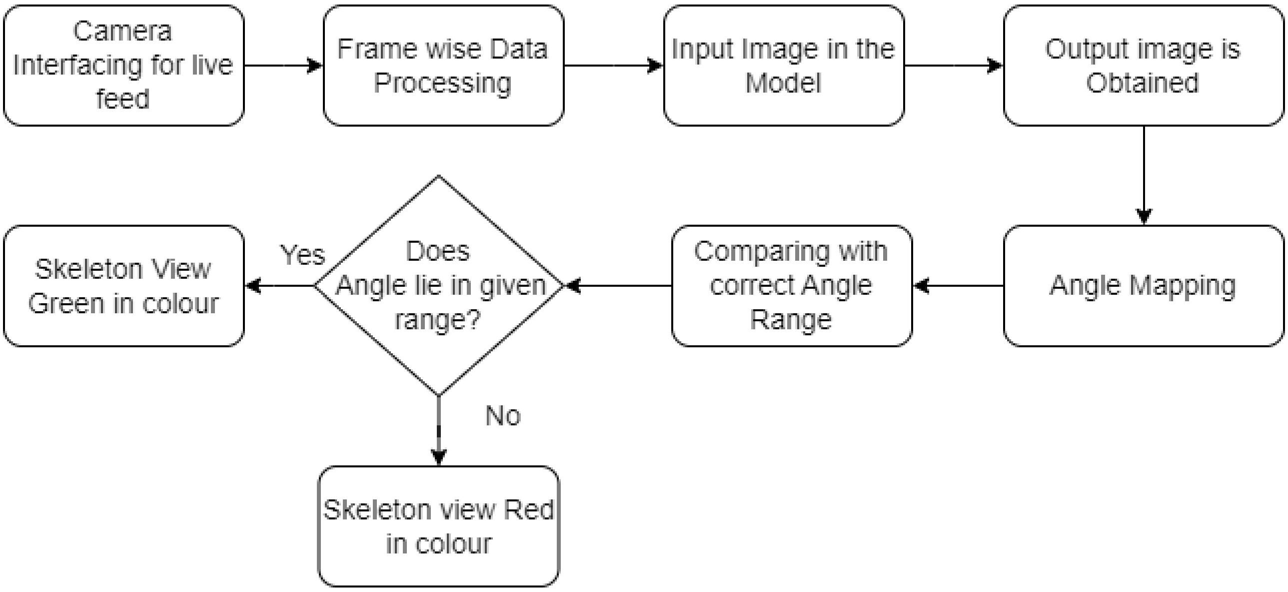

Figure 7: Flowchart of the procedure

Step 1: Camera Interface for Real-Time Feed: To record a live video feed, connect the camera to the system.

Step 2: Input of Frame-wise Data: In order to collect individual frames as input data for analysis, process the live feed frame by frame.

Step3: Image Input for the Model: Put the picture data into the pose detection model for every frame. Frameworks like MediaPipe, which may approximate important points and the frame's skeleton, can serve as the foundation for this concept.

Step 4: The desired image is obtained: Obtain the pose detection model's output image. Visual representations of the identified skeletal structure, important details, and any other pertinent data may be included in this graphic.

Step 5: Measurement of Angles: To determine the angles between particular joints or body parts, analyze the generated image. This can entail making geometrical computations using the coordinates of important points the model identified.

Step 6: Comparing utilizing the proper angle range: Check that the measured angles match the appropriate range of angles for the particular yoga pose being done. This range of acceptable angles might be predetermined by anatomical principles, expert knowledge, or yoga rules.

Step 7: To verify that the measured angles are within the specified range that is appropriate for the yoga pose.

Step 8: Choose a color scheme to represent the skeletal structure if the angles are within the proper range. For instance, to show proper alignment, make the skeleton appears green; if not, to show misalignment or improper posture, display the skeletal structure in a different hue, i.e., red if the angles are outside of the acceptable range.

The developed Machine Learning model is intended to be deployed on the web that can be accessed by users through the internet. The website is designed to respond to the asana that is selected by the user. The section also provides a basic information about the asana. The model is updated and triggered based on the selected asana.

On the initiation of the process, live video feeds are sent frame by frame as the input to the model. The frames are processed and are fed as the input to the model and the body contours are mapped, according to the programmed MediaPipe/MoveNet model, which is the output after which the angle mapping is performed.

We conducted a comparison of the two models and subsequently trained them both in order to determine their respective accuracies. Our analysis ultimately revealed that, although Movenet exhibits exceptional speed in detecting skeleton views, its level of accuracy does not meet the desired standard. Conversely, Mediapipe boasts significantly higher accuracy while maintaining satisfactory detection speeds, rendering it an optimal choice for our specific use case of yoga pose estimation. Consequently, we proceeded with the development of our website utilizing the Mediapipe model.

Table 1: Comparison between Mediapipe and Movenet

Feature	**MediaPipe**	**MoveNet**
Input	Single image or video frame	Single image or video frame
Output	Keypoints, hand landmarks, face landmarks, pose world landmarks	Keypoints
Applications	Augmented reality, virtual reality, gesture recognition, fitness tracking	Pose estimation, rapid action recognition, motion analysis
Accuracy	High	Low
Latency	Low	Low
Model size	Large	Small
Computational cost	High	Low
Ease of use	Yes	No
Community support	Large	Small
Open source	Yes	Yes

Angles for individual models are defined based on the correct way of performing that asana. When the user performs the selected asana, the contours of the user's body are mapped and the significant angles are mapped. If the mapped angles lie within the range of the defined correct angles for that asana, the user is notified that the performed asana is correct by displaying green mappings, otherwise red for the wrong exhibition.

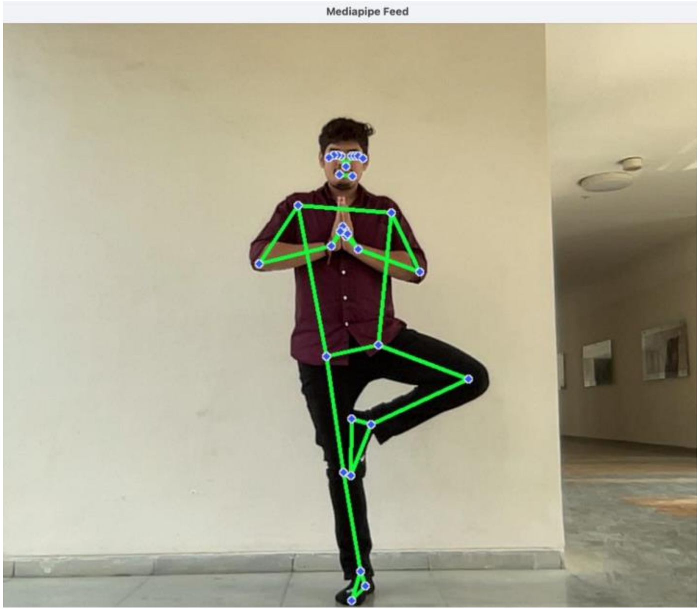

Figure 8: Correct posture for Tree pose mapped in green color

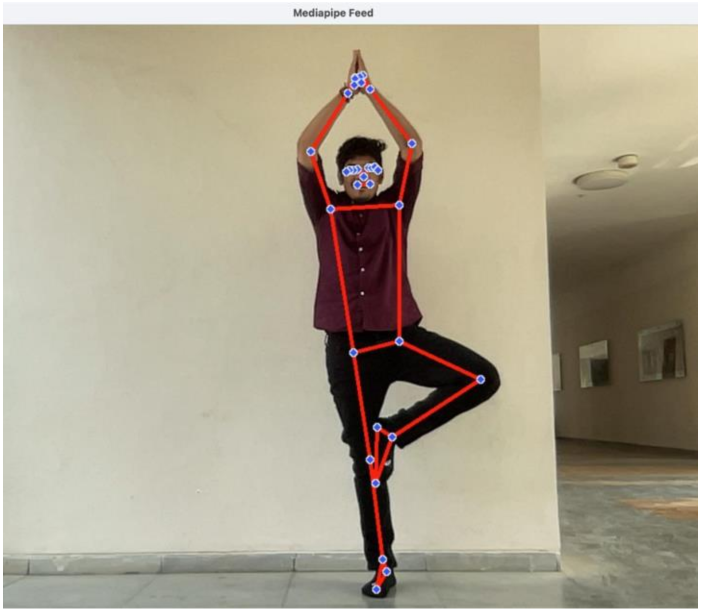

Figure 9: Incorrect posture for Tree pose mapped in red color

Detailed information about the model and the deployment along with the user's trial were described in the aforementioned sections of the report.

Chapter 5: RESULT, ANALYSIS AND DISCUSSIONS

The mapped contours of the five selected yoga asanas (Downdog, Goddess, Tree, Plank & Warrior) were generated using both MediaPipe and MoveNet. Visual comparisons reveal that both models were successfully able to capture the key points of each pose, with slight variations in contour details. MediaPipe demonstrated a better understanding of specific body movements, resulting in a more accurate representation of yoga poses in comparison to MoveNet.

Localization Results:

1. Adho Mukha Svanasana (Downdog)

Figure 10. Mapped contours of DownDog Pose

2. Utkata Konasana (Goddess)

Figure 11. Mapped contours of Goddess Pose

3. Vrikshasana (Tree)

Figure 12. Mapped contours of Tree Pose

4. Phalakasana

Figure 13. Mapped contours of Plank Pose

5. Virabhadrasana (Warrior)

Figure 14. Mapped contours of Warrior Pose

Table 2: Asanas & Model Accuracies

Asana	**Mediapipe [1]**			**MoveNet[9]**		
	No. of correct postures [Out of(Total images)]	**No. of Predicted correct postures**	**Accuracy (%)**	**No. of correct postures [Out of (Total images)]**	**No. of Predicted correct postures**	**Accuracy (%)**
Adho Mukha Svanasana	77(103)	76	98.7	77(103)	70	90.9
Utkata Konasana	42(59)	42	100	42(59)	37	88.09
Vrikshasana	23(38)	23	100	23(38)	18	78.26
Phalakasana	49(66)	49	100	49(66)	42	85.71
Virabhadrasana	91(124)	91	100	91(124)	80	87.91

Quantitative evaluation of accuracy was performed by comparing the number of correct postures identified by each mode using the data collected in Table. 1. It was observed that MediaPipe exhibited a higher accuracy, correctly predicting a larger proportion of yoga poses in the dataset when compared to MoveNet.

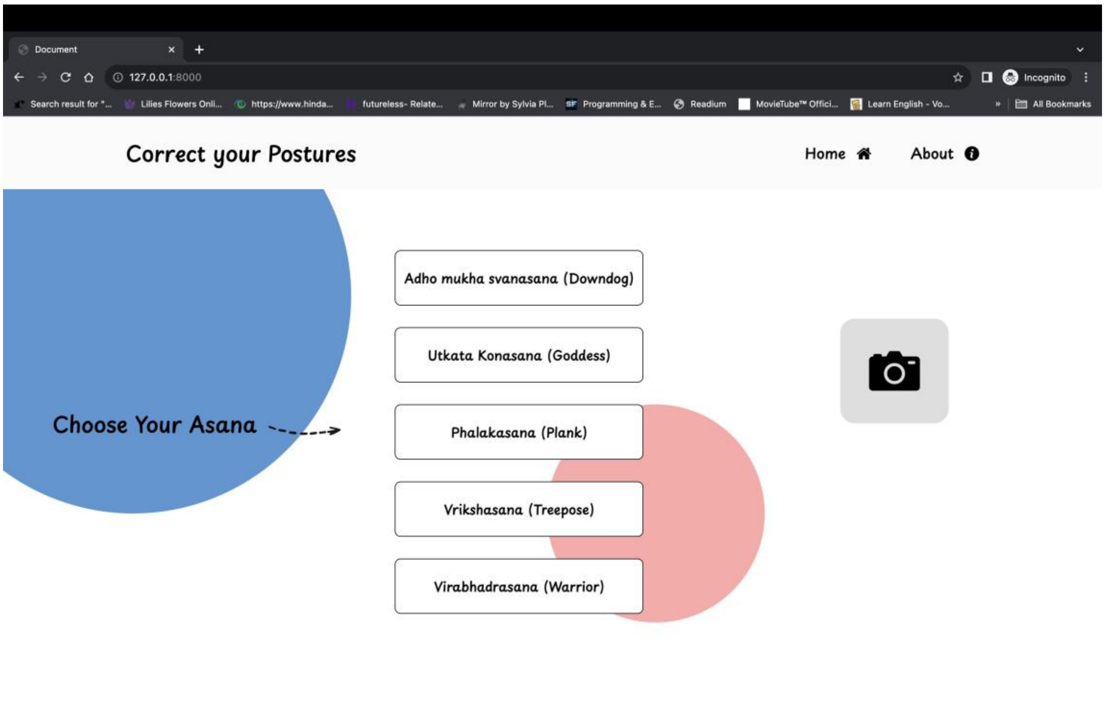

Figure 15: User Interface of the website

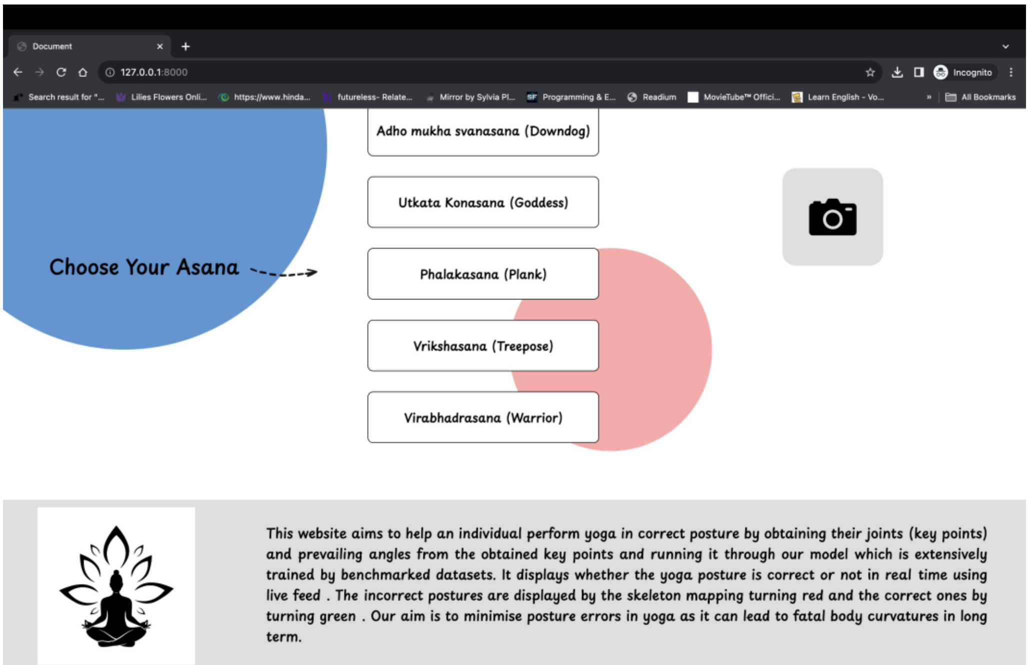

Figure 16: About section of the website

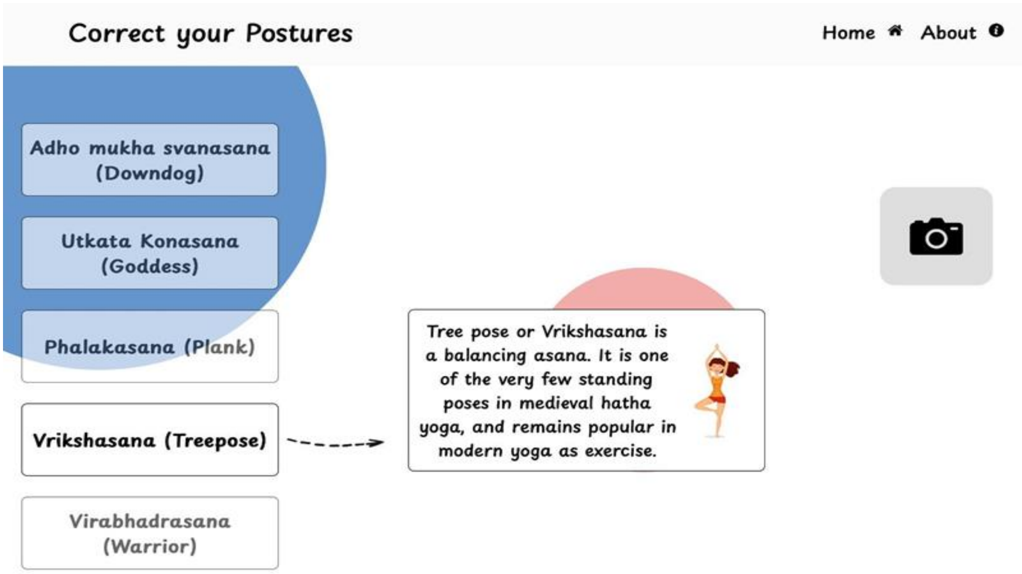

Figure 17: Interface after the user selects an asana Vrikshasana

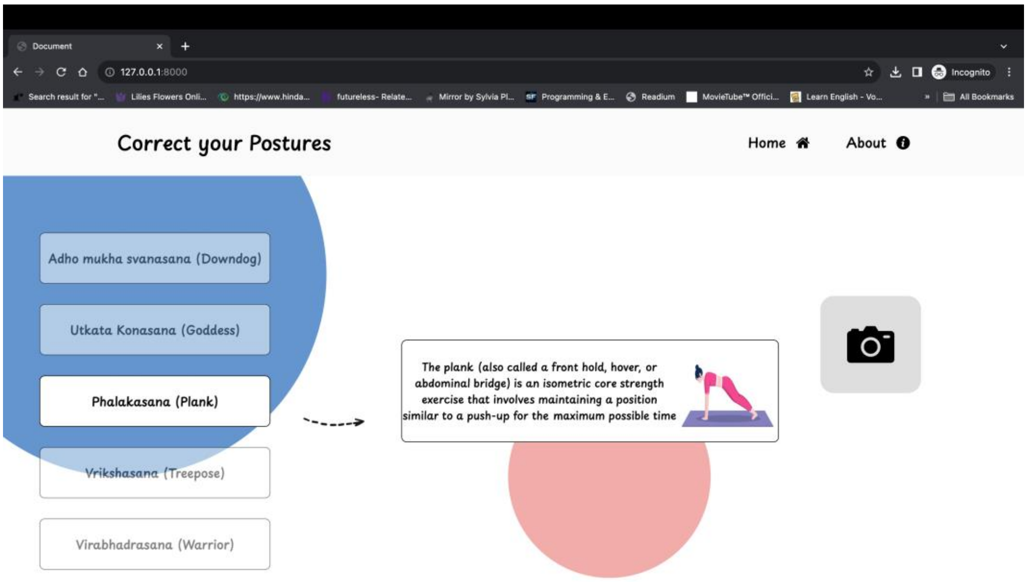

Figure 18: Interface after the user selects an asana Phalakasana.

Our user-friendly website offers an immersive experience for yoga enthusiasts. Users can choose an asana and perform their postures in real-time in front of a webcam, receiving instant feedback powered by our pose estimation models. Users also get information about the asanas and the way to perform it if they are new to that asana.

Visual Results:

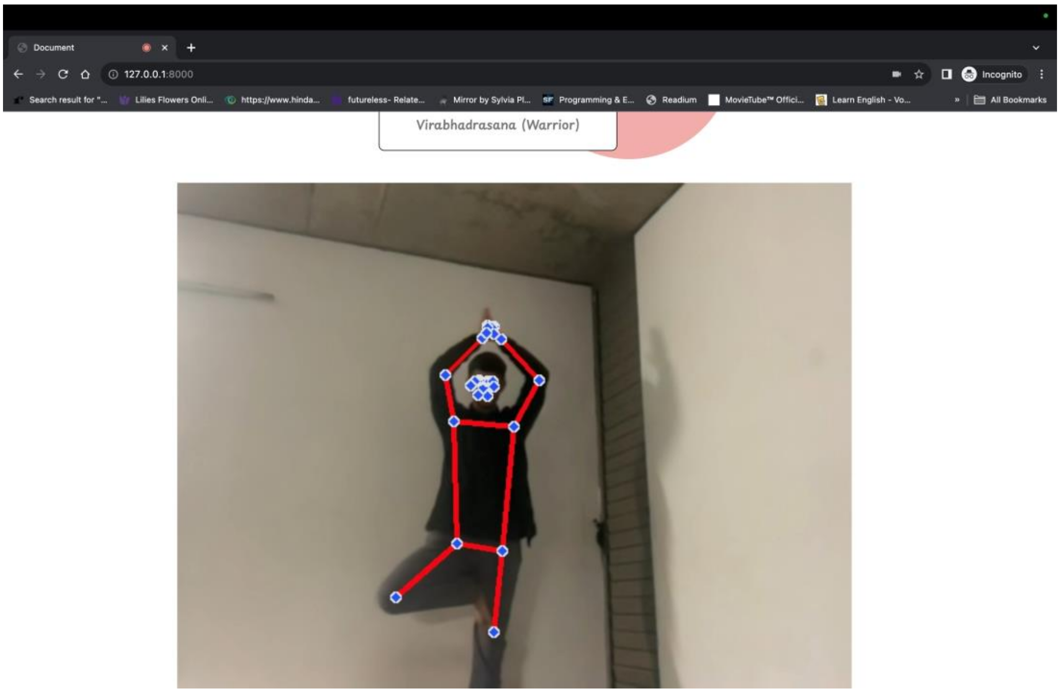

Figure 19: Live feed for incorrect posture for tree pose being performed

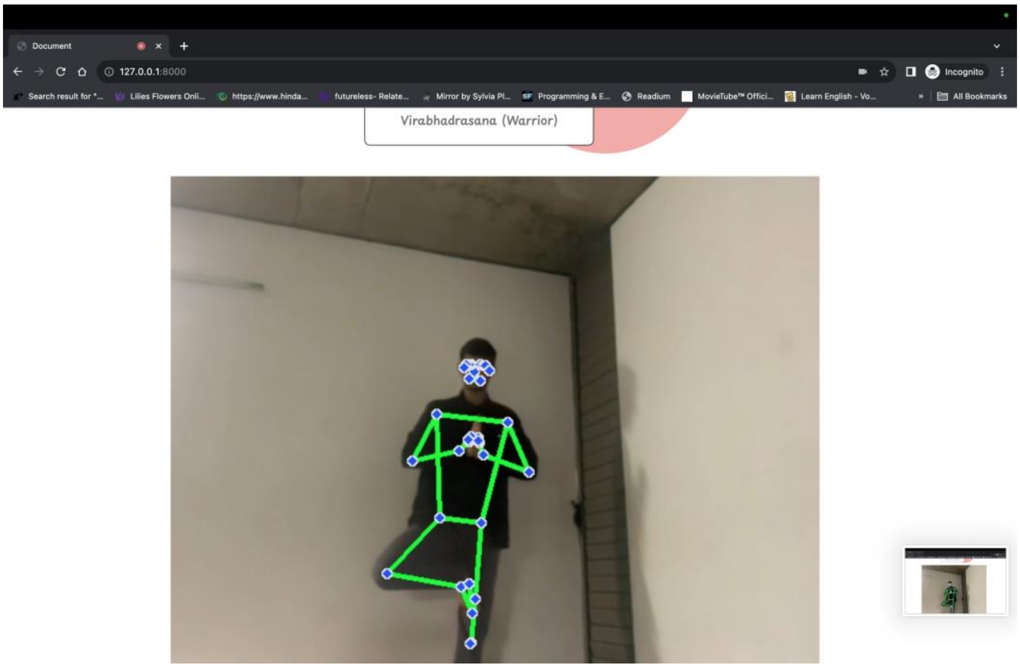

Figure 20: Live feed for correct posture for tree pose being performed

Figure 21: Live feed for incorrect posture for Goddess pose being performed

Figure 22: Live feed for correct posture for Goddess pose being performed

Real-time evaluation using MediaPipe involved performing a yoga asana in front of a webcam to assess the model's ability to provide live feedback. The model showcased remarkable accuracy in capturing and mapping the key points in the body in real time and providing feedback on whether the asana performed is correct or needs to be changed. Fig. 12 shows a green marking on the user input showing that the asana performed has a

correct posture whereas Fig. 13 shows a red marking on the user input showing that the asana performed has an incorrect posture. This feature enhances the tool's usability, allowing users to receive immediate feedback on their posture during a live practice session.

CONCLUSION

In this report, we propose MediaPipe and MoveNet, two robust pose estimation models, in the context of yoga. Our focus centered on five distinct yoga asanas (Downdog, Goddess, Tree, Plank & Warrior), each selected from a comprehensive dataset, forming the foundation for our investigation.

Through the utilization of MediaPipe and MoveNet, we sought to unravel the efficacy of these models in accurately predicting and comparing the output of yoga poses. The models were put to the test, leveraging their unique architectures and real-time processing capabilities, to provide insights into their performance on a diverse set of yoga postures.

The comparison of the model outputs and the ground truth from the dataset allowed us to gauge the accuracy and reliability of MediaPipe and MoveNet in the specific domain of yoga pose estimation. The results unveiled the strengths and nuances of each model, shedding light on their effectiveness in capturing the intricacies of various yoga asanas.

After comparing both models, MediaPipe and MoveNet, for predicting yoga poses, we found that MediaPipe was more accurate. Because of this, we've decided to use MediaPipe in our yoga pose correction website. This choice is based on the fact that MediaPipe performed better, giving us confidence that it will help users correct their yoga poses more effectively.

We believe that by using MediaPipe, we can offer users a better experience on our website, helping them improve their yoga practice. This decision reflects our commitment to using the right technology to make our website user-friendly and supportive for people working on their well-being.

FUTURE SCOPE

While this report provides valuable insights into the use of Computer Vision models for yoga pose estimation, there are several ways in which it can be enhanced by future exploration:

- Further improvement of the MediaPipe model specifically for yoga poses could potentially improve its accuracy. Optimization techniques, such as model compression or quantization, may also be explored to make the model more efficient. This can be more helpful to users practicing yoga.

- Expanding the dataset to include a wider variety of yoga asanas would contribute to the model's robustness. A more diverse dataset could lead to improved generalization to real-world scenarios.

- Enhancing the real-time feedback mechanism on the website could make it more interactive and engaging.

- We have encountered some latency in the website in the sending and fetching of the data frames. This leads to less real-time functioning of the developed machine learning model. We aim to reduce this latency to the least possible.

- We are considering the development of a mobile application that integrates the pose correction tool to attract users who prefer practicing yoga on their mobile devices. This could expand the reach of the tool to a wider audience.

REFERENCES

[1]. C. Lugaresi et al., "MediaPipe: A Framework for Building Perception Pipelines," arXiv:1906.08172[cs.DC].2019.

[2]. M. C. Thar, K. Z. N. Winn, and N. Funabiki, "A Proposal of Yoga Pose Assessment Method Using Pose Detection for Self-Learning," 2019 International Conference on Advanced Information Technologies (ICAIT), Yangon, Myanmar, 2019, pp. 137-142, doi: 10.1109/AITC.2019.8920892.

[3]. M. Chasmai, N. Das, A. Bhardwaj, and R. Garg, "A View Independent Classification Framework for Yoga Postures," SN Computer Science, vol. 3, no. 6, Sep. 2022, doi: https://doi.org/10.1007/s42979-022-01376-7.

[4]. M. Desai and H. Mewada, "A novel approach for yoga pose estimation based on in-depth analysis of human body joint detection accuracy," PeerJ Computer Science, vol. 9, p. e1152, Jan. 2023, doi: https://doi.org/10.7717/peerj-cs.1152.

[5]. S. Garg, A. Saxena, and R. Gupta, "Yoga pose classification: A CNN and MediaPipe inspired deep learning," June 2022 Journal of Ambient Intelligence and Humanized Computing, doi:10.1007/s12652-022-03910-0.

[6]. S. K. Yadav, A. Singh, A. Gupta, and J. L. Raheja, "Real-time Yoga recognition using deep learning," Neural Computing and Applications, May 2019, doi: https://doi.org/10.1007/s00521-019-04232-7.

[7]. V. Anand Thoutam et al., "Yoga Pose Estimation and Feedback Generation Using Deep Learning," Computational Intelligence and Neuroscience, vol. 2022, pp. 1–12, Mar. 2022, doi: https://doi.org/10.1155/2022/4311350.

[8]. Y. Suryawanshi, N. Gunjal, B. Kanorewala, and K. Patil, "Yoga dataset: A resource for computer vision-based analysis of Yoga asanas," Data in Brief, vol. 48, p. 109257, Jun. 2023, doi: https://doi.org/10.1016/j.dib.2023.109257.

[9] “MoveNet: Ultra fast and accurate pose detection model.,” *TensorFlow*. https://www.tensorflow.org/hub/tutorials/movenet

[10] “VectorStock - Vector Art, Images, Graphics & Clipart,” *VectorStock*. https://www.vectorstock.com/

Final_draft

ORIGINALITY REPORT

9%	7%	8%	3%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	downloads.hindawi.com Internet Source	2%
2	www.ncbi.nlm.nih.gov Internet Source	1%
3	link.springer.com Internet Source	1%
4	Maybel Chan Thar, Khine Zar Ne Winn, Nobuo Funabiki. "A Proposal of Yoga Pose Assessment Method Using Pose Detection for Self-Learning", 2019 International Conference on Advanced Information Technologies (ICAIT), 2019 Publication	1%
5	www.researchgate.net Internet Source	1%
6	www.ijeast.com Internet Source	<1%
7	Submitted to University of Information Technology, Yangon Student Paper	<1%

8	Ranjana Jadhav, Vaidehi Ligde, Rushikesh Malpani, Phinehas Mane, Soham Borkar. "Aasna: Kinematic Yoga Posture Detection And Correction System Using CNN", ITM Web of Conferences, 2023 Publication	<1 %
9	Santosh Kumar Yadav, Amitojdeep Singh, Abhishek Gupta, Jagdish Lal Raheja. "Real-time Yoga recognition using deep learning", Neural Computing and Applications, 2019 Publication	<1 %
10	Submitted to Saint Louis University Student Paper	<1 %
11	peerj.com Internet Source	<1 %
12	www.mdpi.com Internet Source	<1 %
13	www.scilit.net Internet Source	<1 %
14	Submitted to Indian Institute of Information Technology Sri City Student Paper	<1 %
15	Submitted to Temple University Student Paper	<1 %
16	Submitted to Liverpool John Moores University Student Paper	<1 %

17	Vivek Anand Thoutam, Anugrah Srivastava, Tapas Badal, Vipul Kumar Mishra et al. "Yoga Pose Estimation and Feedback Generation Using Deep Learning", Computational Intelligence and Neuroscience, 2022 Publication		<1 %
18	ijarsct.co.in Internet Source		<1 %
19	Submitted to CSU, San Jose State University Student Paper		<1 %
20	lutpub.lut.fi Internet Source		<1 %

Exclude quotes	On	Exclude matches	< 11 words
Exclude bibliography	On		