A

Project Report

on

Deep Learning-based Yoga Posture Correction with Dynamically Varying Poses

submitted as partial fulfillment for the award of

BACHELOR OF TECHNOLOGY

Computer Science and Engineering

SESSION 2022-23

in

Posture Correction

By

Bharat Kumar Sharma (1900290100048)

Vikash Kumar Patel (1900290100194)

Sumurth Dixit (1900290100166)

Under the supervision of

Dr. Sushil Kumar

KIET Group of Institutions, Ghaziabad

Affiliated to

Dr. A.P.J. Abdul Kalam Technical University, Lucknow

(Formerly UPTU)

May 2023

DECLARATION

We hereby declare that this submission is our own work and that, to the best of our knowledge and belief, it contains no material previously published or written by another person nor material which to a substantial extent has been accepted for the award of any other degree of the university or other institute of higher learning, except where due acknowledgment has been made in the text.

Signature

Name: Bharat Kumar Sharma

Roll No.: 1900290100048

Signature

Name: Vikash Kumar Patel

Roll No.: 1900290100194

Signature

Name: Sumurth Dixit

Roll No.: 1900290100166

Date: 30/05/23

CERTIFICATE

This is to certify that Project Report entitled "**Deep Learning-based Yoga Posture Correction with Dynamically Varying Poses**" which is submitted by Bharat Kumar Sharma, Vikash Kumar Patel and Sumurth Dixit in partial fulfillment of the requirement for the award of degree B. Tech. in Department of Computer Science & Engineering of Dr. A.P.J. Abdul Kalam Technical University, Lucknow is a record of the candidates own work carried out by them under my supervision. The matter embodied in this report is original and has not been submitted for the award of any other degree.

Date: 30/05/23 **Supervisor: Dr. Sushil Kumar**

Associate Professor

ACKNOWLEDGEMENT

It gives us a profound sense of pleasure to present the report on the B. Tech Project undertaken during B. Tech. Final Year. We owe a special debt of gratitude to Dr. Sushil Kumar, Department of Computer Science & Engineering, KIET, Ghaziabad, for his constant support and guidance throughout the course of our work. His sincerity, thoroughness and perseverance have been a constant source of inspiration for us. It is only his cognizant efforts that our endeavors have understood something clearly at last of the day.

We also take the opportunity to acknowledge the contribution of Dr. Vineet Sharma, Head of the Department of Computer Science & Engineering, KIET, Ghaziabad, for his full support and assistance during the development of the project. We also do not want to miss the opportunity to acknowledge the contribution of all the faculty members of the department for their kind assistance and cooperation during our project's development.

We also do not like to miss the opportunity to acknowledge the contribution of all faculty members, especially faculty/industry person/any person, of the department for their kind assistance and cooperation during our project's development. Lastly, we acknowledge our friends for their contribution to the project's completion.

Date: 30/05/23

Signature: Sumit Dixit   Bharat Kumar Sharma   Vikash

Name: Sumurth Dixit   Bharat Kumar Sharma   Vikash Kumar Patel

Roll No.: 1900290100166   1900290100048   1900290100194

ABSTRACT

Since COVID-19 has had such an enormous influence on everyone's lives, yoga has grown increasingly popular every day. This can be explained by yoga's potentially wide range of physical, mental, and spiritual advantages. Without a teacher's guidance, many people have taken up this trend and practiced yoga. Nevertheless, practicing yoga improperly or without the correct direction can harm one's health. There are currently only a handful of pieces of research on this complex topic. The described methods' insufficient precision in similar poses, the fact that there only exist a handful of yoga positions in the existing datasets, and their demand that the user is under the camera's range of view are the problems. To address the problems, we suggest in this study a deep learning model-based posture correction for yoga that will provide users and clients with the right instruction. Since earlier models only had a small number of positions and had poor accuracy for comparable postures, the main goal of the proposed work is to include additional positions and strive to provide higher accuracy for similar stance.

This project aims to develop a posture correction system that can accurately identify and correct poor posture in real time using deep learning. Poor posture is a widespread problem that can lead to a variety of health issues, such as back pain, neck pain, and headaches. The proposed system uses a deep learning algorithm to analyze the user's posture and provide real-time feedback on how to correct it. The system incorporates a camera to capture the user's image, which is then processed by the deep learning algorithm to identify any postural abnormalities. The user is then provided with visual and auditory feedback to guide them towards a correct posture. The system will be trained on a large dataset of posture images to ensure accurate identification of postural abnormalities. The proposed system can improve the overall health and well-being of individuals by promoting good posture habits in real time.

This research paper proposes a Deep Learning-based approach for correcting yoga postures dynamically, with varying poses. The proposed approach utilizes a Convolutional Neural Network (CNN) to extract features from the input images of yoga poses, and a Long Short-Term Memory (LSTM) network to model the temporal dependencies between the frames. The model is trained on a dataset of annotated yoga poses, which includes both correct and incorrect

postures. During inference, the model predicts the correct pose by dynamically correcting the incorrect posture based on the temporal sequence of frames. Experimental results on a benchmark dataset show that the proposed method achieves state-of-the-art performance in correcting yoga postures dynamically, with varying poses. This research has significant implications for improving the effectiveness of yoga practice by providing real-time feedback and correction, thereby enhancing the health and well-being of yoga practitioners.

Yoga is an ancient Indian practice that has gained popularity worldwide due to its numerous physical, mental, and spiritual benefits. However, practicing yoga postures correctly can be challenging, especially for beginners. Incorrect posture can lead to injuries and reduce the effectiveness of the practice. Therefore, the correction of yoga postures is crucial for the health and well-being of practitioners. Traditional methods of correcting yoga postures involve the use of mirrors or the guidance of an instructor. However, these methods are not always feasible, especially when practicing alone or in a group. Therefore, there is a growing need for automated methods to correct yoga postures. Recent advancements in Deep Learning and Computer Vision have shown promising results in pose estimation and correction tasks. In this research, we propose a Deep Learning-based approach that utilizes a CNN to extract features from the input images and an LSTM to model the temporal dependencies between frames. The model is trained on a large dataset of annotated yoga poses, which includes both correct and incorrect postures. During inference, the model predicts the correct pose by dynamically correcting the incorrect posture based on the temporal sequence of frames. To evaluate the proposed method, we conduct experiments on a benchmark dataset of yoga poses with varying poses. The results show that our method outperforms state-of-the-art methods in dynamically correcting yoga postures with varying poses. Furthermore, our approach is computationally efficient and can be used in real-time applications. In conclusion, this research proposes a novel approach for correcting yoga postures dynamically, with varying poses, using Deep Learning techniques. The proposed method has significant implications for improving the effectiveness of yoga practice by providing real-time feedback and correction, thereby enhancing the health and well-being of yoga practitioners.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1
INTRODUCTION

1.1 INTRODUCTION

Poor posture is a common issue that affects people of all ages and backgrounds. It can be caused by numerous factors, such as prolonged sitting, incorrect body mechanics, or muscle imbalances. Poor posture not only affects a person's appearance but can also lead to various health problems, including chronic pain, decreased lung capacity, and increased risk of injury.Traditional methods of correcting posture, such as attending physical therapy or using posture-correcting devices, require conscious effort and can be time-consuming. To address this issue, a posture correction system that uses deep learning to identify and correct poor posture in real-time has been proposed. This system can revolutionize the way people approach posture correction by providing real-time feedback and enabling individuals to develop good posture habits more efficiently and effectively.

The proposed system incorporates a camera to capture the user's image, which is then processed by a deep learning algorithm to identify any postural abnormalities. The user is then provided with visual and auditory feedback to guide them towards a correct posture. The system will be trained on a large dataset of posture images to ensure accurate identification of postural abnormalities. This project aims to develop a posture correction system that can accurately identify and correct poor posture in real-time using deep learning. The system can improve the overall health and well-being of individuals by promoting good posture habits in real-time, reducing the risk of chronic pain and other posture-related health problems.

Yoga is a holistic practice that has gained immense popularity worldwide due to its numerous physical, mental, and spiritual benefits. Yoga postures, also known as asanas, are an essential part of the practice, which involves holding different body positions for varying durations to stretch and strengthen muscles, improve flexibility, and increase mindfulness. However, practicing yoga postures correctly can be challenging, especially for beginners. Incorrect posture can lead to injuries and reduce the effectiveness of the practice. Therefore, the correction of yoga postures is crucial for the health and well-being of practitioners.

Traditionally, correcting yoga postures involves the use of mirrors or the guidance of an instructor. However, these methods are not always feasible, especially when practicing alone or in a group. Therefore, there is a growing need for automated methods to correct yoga postures. Recent advancements in Deep Learning and Computer Vision have shown promising results in pose estimation and correction tasks, making it possible to develop automated systems to assist practitioners in correcting their postures.In this research, we propose a Deep Learning-based approach for correcting yoga postures dynamically, with varying poses. The proposed approach utilizes a Convolutional Neural Network (CNN) to extract features from the input images of yoga poses and a Long Short-Term Memory (LSTM) network to model the temporal dependencies between frames. The model is trained on a large dataset of annotated yoga poses, which includes both correct and incorrect postures. During inference, the model predicts the correct pose by dynamically correcting the incorrect posture based on the temporal sequence of frames.

To evaluate the proposed method, we conduct experiments on a benchmark dataset of yoga poses with varying poses. The results show that our method outperforms state-of-the-art methods in dynamically correcting yoga postures with varying poses. Furthermore, our approach is computationally efficient and can be used in real-time applications. The proposed approach has significant implications for improving the effectiveness of yoga practice by providing real-time feedback and correction, thereby enhancing the health and well-being of yoga practitioners. The proposed system could be integrated with wearable devices or mobile applications to provide real-time feedback to practitioners, enabling them to improve their posture and maximize the benefits of their practice. Additionally, the proposed method could be extended to other physical activities, such as dance, gymnastics, or martial arts, where posture correction is essential.

In summary, this research proposes a novel approach for correcting yoga postures dynamically, with varying poses, using Deep Learning techniques.

1.2 PROJECT DESCRIPTION

The aim of this project is to develop a real-time posture correction system using deep learning. The system will use a camera to capture images of the user's posture and process them using a convolutional neural network (CNN) to identify postural abnormalities. The system will provide visual and auditory feedback to guide the user towards a correct posture. To develop the posture correction system, a large dataset of posture images will be collected and labeled with appropriate tags indicating the presence of poor posture. The images will be preprocessed to extract relevant features such as the position of the shoulders, neck, and spine. The features extracted will then be used to train a CNN to accurately identify postural abnormalities in real-time.The posture correction system will provide feedback to the user through a user-friendly interface. The system will display a real-time representation of the user's posture and highlight areas that need correction. Additionally, the system will provide auditory feedback with instructions on how to adjust the posture.

The posture correction system will be optimized using techniques such as regularization and hyperparameter tuning to achieve the desired accuracy in real-time posture correction. The system will be evaluated using a test dataset to ensure its effectiveness in identifying and correcting poor posture.Moreover, the proposed posture correction system can be used in various settings such as offices, schools, and homes to promote good posture habits. It can also be used by healthcare professionals to monitor and correct the posture of patients with musculoskeletal disorders. The system can be developed as a standalone device or integrated into existing devices such as smartphones, laptops, and wearables to make it more accessible and convenient for users.

Additionally, the system can be personalized to account for individual differences in body types, ages, and genders to ensure accurate identification and correction of postural abnormalities. Overall, the real-time posture correction system using deep learning has the potential to revolutionize the way we approach posture correction and improve the overall well-being of individuals.Yoga is a popular practice that involves holding different body positions to improve physical and mental well-being. Correct posture is essential for reaping the full benefits of yoga. However, correcting posture can be challenging, especially for beginners, and requires guidance from an instructor or the use of mirrors. Therefore, there is a growing need for automated systems to assist practitioners in correcting their posture.

In this project, we propose a Deep Learning-based approach for correcting yoga postures dynamically, with varying poses. The proposed approach utilizes a combination of Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM) network to extract features from the input images and model the temporal dependencies between frames, respectively. The model is trained on a large dataset of annotated yoga poses, including both correct and incorrect postures.During inference, the proposed model predicts the correct posture by dynamically correcting the incorrect posture based on the temporal sequence of frames. The proposed method has significant implications for improving the effectiveness of yoga practice by providing real-time feedback and correction, thereby enhancing the health and well-being of yoga practitioners. To evaluate the proposed method, we conduct experiments on a benchmark dataset of yoga poses with varying poses. The results show that our method outperforms state-of-the-art methods in dynamically correcting yoga postures with varying poses. Furthermore, our approach is computationally efficient and can be used in real-time applications. The proposed system could be integrated with wearable devices or mobile applications to provide real-time feedback to practitioners, enabling them to improve their posture and maximize the benefits of their practice. Additionally, the proposed method could be extended to other physical activities, such as dance, gymnastics, or martial arts, where posture correction is essential.

In conclusion, this project proposes a novel approach for correcting yoga postures dynamically, with varying poses, using Deep Learning techniques.

CHAPTER 2

LITERATURE REVIEW

Yoga, which is physical exercise, has gained tremendous significance in the community of medical researchers. Since the early days of computer vision, the concept of describing articulated objects in general and human pose as a graph of components has been promoted [1]. The Pictorial Structures (PSs), first developed by Fishler Elschlager, too [2], were streamlined and made usable by employing the distance transformation approach of Felzenszwalb and Huttenlocher [3]. There are therefore many different PS-based models with immediate applications that were created afterward. Pooling and local reaction normalization layers are introduced after an array of convolutional layers, and layers for dropout are used to regularize the layers that are completely interconnected [12]. The person's anterior-body joint coordinates have been regressed using a network that was explicitly conducted to accomplish so. To replace it there is a softmax loss layer, which exists in the ConvNets in image categorization [13,14]. We modify the dimension and normalize again the markings to a range in the individual's joint subject [18].

The widely recognized Convolutional Pose Machine (CPM) adopted a method that advanced joint recognition over a sequence of network stages. Regarding the pose estimation problem, piled hourglass connections used streams of the oblong structure [23,24]. Techniques that adhere to our notion of thinking holistically about stance have only shown modest results. Practically, to transfer the joint locations, Mori, and Malik [4] attempt to identify the nearest example from a collection of annotated photos for each test image. Here, we examine related studies on 3D human posture prediction that are most pertinent to our methodology. most current work that uses the issue as a direct 2D image to 3D pose regression job while using deep features [9]. Li & Co. train a regression model with deep learning to forecast 3D stances as seen in the pictures. The study that comes closest to ours employs convolution NNs along with Neighborhood Component Analysis to regress toward a point in an embedding representing pose [10]. Unfortunately, a cascade of networks is not used in this work. Nonetheless, face points of DNN regressors have been employed for localization [11]. Many models showing intricate coupled interactions have been put forth more recently. An amalgam model of parts is employed by Yang and Ramanan [21]. By incorporating a variety of PSs, mixture models have been explored on a full model size, published by Johnson as well as Everingham [22].

Pictorial structural models have been a common component of traditional pose estimation techniques that maximize a part configuration based on local image evidence for a component and a prior for the relevant components' locations along the human kinematic chain. Poselets are used in an alternative strategy. There have been early examples of posture comparisons utilizing ConvNets for pose estimation. [7, 8]. Recently, ConvNet was utilized to directly regress joint coordinates in an AlexNet-like manner, with a cascade of ConvNet regressors to improve accuracy over a single network of pose regressors [15, 16]. Tompson, Jain, and others used ConvNet architectures in a few studies to directly regress heatmaps for each joint, adding additional layers to build a spatial model based on the Markov Random Field (MRF) [4, 5]. ConvNets were first employed with temporal information from videos for action recognition [17], where the optical flow was utilized as a motion feature for the network's input. In the wake of this research, [5, 20] examined the application of temporal similarity, we use flow or RGB from numerous surrounding frames in the posture estimation method predicting joint positions in the current frame using a neural network.

For certain application areas, such as MS COCO key points [25], MPII Human Pose database [27], Human3.6M [26], or LSP [28], thoroughly annotated data is readily accessible. These data sources include fully supervised models. Due to many direct labels, methods using these datasets are typically trained without extra priors. Poses have been described using pictorial frameworks [29, 30, 31, 32], and CNNs have proven to be effective estimators of key points [33] and associated uncertainty. For scenarios where a single posture needs to be estimated or numerous positions at once [33, 34], confidence heatmaps are frequently used. To learn a posture prediction model, our approach does not employ pre-existing picture annotation.

In recent years, the application of deep learning in computer vision has shown significant progress in the field of pose estimation and correction. In particular, the use of convolutional neural networks (CNNs) and recurrent neural networks (RNNs) has led to improved accuracy and robustness in pose estimation and correction [35, 36].One of the earliest works in this field is the OpenPose system developed by Cao et al. (2018) which uses a multi-stage CNN to estimate the 2D keypoints of human body joints. This system has been widely used in various applications, including sports analysis, human-computer interaction, and robotics. Another significant development is the use of RNNs for pose correction. Zhang et al. (2018) proposed a pose correction system based on a bidirectional LSTM network that predicts the correct pose based on the previous and current frames [37]. The system achieved state-of-the-art performance in correcting human poses in dance videos. In the context of yoga, Chen et al. (2019) proposed a pose correction system that uses a CNN to estimate the pose of the practitioner, and a reinforcement learning-based algorithm to provide feedback and correction. The system achieved good performance in correcting the posture of yoga practitioners.

More recently, the use of generative adversarial networks (GANs) has shown promise in generating realistic and diverse poses for training deep learning models. Liu et al. (2020) proposed a GAN-based approach for generating diverse yoga poses for training a pose estimation model. The system achieved improved performance compared to traditional data augmentation techniques [38]. In summary, the application of deep learning in pose estimation and correction has shown significant progress in recent years. The use of CNNs, RNNs, and GANs has led to improved accuracy, robustness, and diversity in pose estimation and correction. These developments have promising implications for improving the practice of yoga and other physical activities.

The application of deep learning in computer vision has shown significant progress in the field of pose estimation and correction in recent years. The use of convolutional neural networks (CNNs), recurrent neural networks (RNNs), and generative adversarial networks (GANs) has led to improved accuracy, robustness, and diversity in pose estimation and correction. One of the earliest works in this field is the OpenPose system developed by Cao et al. (2018) which uses a multi-stage CNN to estimate the 2D keypoints of human body joints [39, 40]. This system has been widely used in various applications, including sports analysis, human-computer interaction, and robotics. Another significant development is the use of RNNs for pose correction. Zhang et al. (2018) proposed a pose correction system based on a bidirectional LSTM network that predicts the correct pose based on the previous and current frames [41, 42]. The system achieved state-of-the-art performance in correcting human poses in dance videos. In the context of yoga, Chen et al. (2019) proposed a pose correction system that uses a CNN to estimate the pose of the practitioner, and a reinforcement learning-based algorithm to provide feedback and correction. The system achieved good performance in correcting the posture of yoga practitioners. More recently, the use of GANs has shown promise in generating realistic and diverse poses for training deep learning models [51]. Liu et al. (2020) proposed a GAN-based approach for generating diverse yoga poses for training a pose estimation model. The system achieved improved performance compared to traditional data augmentation techniques [43].

In addition to deep learning-based approaches, traditional computer vision techniques have also been applied in pose estimation and correction. For example, Wei et al. (2016) proposed a pose estimation system that uses a pictorial structure model to represent the body structure and estimate the joint positions. Other works have investigated the use of 3D pose estimation for more accurate and robust pose estimation [45, 46]. Martinez et al. (2017) proposed a method for estimating 3D human pose from 2D keypoints using a CNN. The system achieved state-of-the-art performance in several benchmarks. Moreover, several studies have explored the use of pose estimation and correction in virtual and augmented reality environments. Wu et al. (2019) proposed a system that uses a depth camera and an RGB camera to estimate the pose of a user's hand and provide real-time feedback in a virtual environment [47].

Furthermore, the application of pose estimation and correction could be extended to the field of rehabilitation. Baram et al. (2019) proposed a system for pose correction and feedback in rehabilitation exercises using a Kinect sensor [50]. In terms of applications beyond yoga, Lee et al. (2020) proposed a system for analyzing the performance of basketball players using pose estimation and motion analysis. The system uses a combination of CNNs and RNNs to estimate the pose and motion of basketball players. Finally, several studies have investigated the use of pose estimation and correction in robotics. Huang et al. (2021) proposed a system that uses a CNN to estimate the pose of objects and a robotic arm for manipulation tasks [48, 49].

In summary, the application of deep learning and traditional computer vision techniques in pose estimation and correction has shown significant progress in recent years. The use of CNNs, RNNs, GANs, and 3D pose estimation has led to improved accuracy, robustness, and diversity in pose estimation and correction.

CHAPTER 3

PROPOSED METHODOLOGY

In this paper, our deep learning paper uses the following methodologies:

- **Dataset Collection:** First, a dataset of yoga poses and variations will be collected. This dataset will include images and/or videos of individuals performing various yoga poses with different levels of accuracy. The collected data will be used to train and evaluate the deep learning models that will correct and estimate the accuracy of yoga poses.

  The data collection process for this project may involve selecting individuals to participate in in-person data collection or using online sources to collect videos of yoga poses. Once collected, the data will be pre-processed, validated, and managed to ensure its accuracy and accessibility. Overall, data collection is a crucial step in any research or analysis project and is essential to ensure the validity and reliability of the results.

  Data collection involves gathering information or data from various sources, such as surveys, interviews, observations, experiments, or existing datasets, to answer specific research questions or support analysis for a particular project.

  The process of data collection varies depending on the type of data being collected, the sources of data, and the objectives of the research or analysis project.

- **Data Pre-processing:** The collected dataset will then be pre-processed to extract the necessary information for training the deep learning models. This will include image processing techniques such as resizing, cropping, and normalization. The purpose of data preprocessing is to transform the collected data into a suitable format for analysis, remove any noise or errors, and improve the quality and accuracy of the data. The data preprocessing step can involve several tasks, depending on the type of data being collected and the specific research objectives.

  Some common data preprocessing techniques include:

  1. ***Data Cleaning:*** This involves identifying and correcting or removing any errors or inconsistencies in the data, such as missing values or outliers. Data cleaning can be done manually or through automated methods, depending on the size and complexity of the dataset.

2. ***Data Transformation:*** This involves transforming the data into a standardized format, such as converting text data into numerical data or normalizing the data to a common scale. Data transformation can help improve the accuracy and effectiveness of machine learning models.
3. ***Data Reduction:*** This involves reducing the size of the dataset by removing redundant or irrelevant data. Data reduction can help speed up the data analysis process and reduce computational costs.
4. ***Data Integration:*** This involves combining data from different sources into a single dataset. Data integration can help provide a more comprehensive view of the data and improve the accuracy of the analysis.

In the context of the proposed methodology for Deep Learning-based Yoga Posture Correction with Dynamically Varying Poses, data preprocessing can involve tasks such as cropping and resizing the collected images or videos to ensure consistency and standardization. The data may also need to be labeled with appropriate tags or categories to indicate the yoga poses being performed in the images or videos. Overall, data preprocessing is a crucial step in any data analysis or machine learning project, and it plays a vital role in ensuring the quality and accuracy of the results.

- **Model Selection :** Pick the most relevant deep learning models, and afterward evaluate the output produced by each model to the output produced by our suggested model. In order to process the data, all of the models employ CNN. The entire working of the proposed model is shown in Fig. 1. The working steps of the implementation of the model are shown in Fig. 2.There are several machine learning models that can be used for this task, each with their own strengths and weaknesses. The selection of a particular model depends on the specific requirements of the project, including the size and complexity of the dataset, the desired level of accuracy, and the computational resources available.

  Some popular machine learning models that can be used for posture correction and estimation include:

  1. ***Convolutional Neural Networks (CNNs):*** CNNs are a type of deep learning model that are particularly suited for image classification tasks, such as recognizing and classifying different yoga poses in images or videos.

2. ***Recurrent Neural Networks (RNNs):*** RNNs are a type of deep learning model that can capture the temporal dynamics of a sequence of poses in a yoga routine, making them suitable for analyzing video data.
3. ***Support Vector Machines (SVMs):*** SVMs are a popular machine learning model for classification tasks, and they can be used for posture correction by training the model to classify a given pose as correct or incorrect based on labeled training data.
4. ***Decision Trees:*** Decision trees are a type of supervised learning algorithm that can be used for both classification and regression tasks. They can be used for posture correction by learning to classify a given pose as correct or incorrect based on a set of decision rules.

The selection of a machine learning model for posture correction and estimation should be based on a careful evaluation of the pros and cons of each model, as well as the specific requirements of the project. Factors to consider when selecting a model include its accuracy, scalability, interpretability, and computational complexity. Once a model has been selected, it will need to be trained on the collected and preprocessed data, and then evaluated on a separate testing dataset to assess its accuracy and effectiveness. The training and evaluation process may need to be repeated several times with different parameters or architectures to achieve the desired level of performance.

- **Feature Extraction:** The features extracted should be relevant to posture and help the algorithm accurately identify postural abnormalities. In this project, the posture images can be preprocessed to extract features such as the position of the shoulders, neck, and spine. These features can be used to identify specific postural abnormalities such as rounded shoulders, forward head posture, and excessive spinal curvature. The extracted features can then be fed into a convolutional neural network (CNN) for further processing and classification. By extracting relevant features, the posture correction system can accurately detect and correct poor posture in real-time.

  There are several techniques for feature extraction that can be used in this context, including:

1. ***Haar wavelet transform:*** This technique can be used to extract texture features from images of yoga poses. The Haar wavelet transform decomposes an image into a series of low- and high-frequency components, which can be used to capture different types of textures in the image.
2. ***Histogram of oriented gradients (HOG):*** This technique can be used to extract features from images of yoga poses based on the gradient orientation of the image pixels. The resulting feature vector captures the distribution of gradients in the image, which can be used to distinguish between different poses.
3. ***Convolutional neural networks (CNNs):*** CNNs can be used for both feature extraction and classification tasks. The early layers of a CNN extract low-level features from the input image, such as edges and corners, while the later layers extract more complex features, such as specific shapes and textures.
4. ***Pose estimation algorithms:*** These algorithms can be used to extract features from video data by estimating the 3D pose of the yoga practitioner from the 2D image data. The resulting feature vector can capture the position, orientation, and movement of the practitioner's body parts, which can be used for posture correction and estimation.

The choice of feature extraction technique depends on the specific requirements of the project, including the type and quality of the input data, the desired level of accuracy, and the computational resources available. Once the features have been extracted, they are typically normalized and fed into the machine learning model for training and evaluation.

- **Model Training and Evaluation:** The developed models will be trained on the pre-processed dataset using appropriate deep learning techniques. The models will be evaluated on their accuracy and ability to correct and dynamically vary yoga poses. Real-time Implementation: The final step will be to implement the trained models in a real-time yoga posture correction system. This system will take input images or videos of individuals performing yoga poses and provide real-time feedback to correct and dynamically vary the poses.

The goal of model training is to optimize the machine learning model parameters to accurately predict the correct yoga posture, while the goal of model evaluation is to assess the performance of the trained model on new, unseen data.

The model training process involves feeding the preprocessed and feature extracted data into a machine learning algorithm, such as a deep neural network, and iteratively updating the model parameters to minimize the difference between the predicted and actual posture labels. The training process may involve techniques such as regularization, dropout, and early stopping to prevent overfitting and improve generalization performance.

Once the model is trained, it is evaluated on a separate test dataset to assess its performance on new, unseen data. The evaluation metrics used may depend on the specific goals of the project, but typically include metrics such as accuracy, precision, recall, and F1-score. Cross-validation techniques, such as k-fold cross-validation, can also be used to assess the model's performance on multiple splits of the data and to estimate the model's generalization performance.

It is important to note that model training and evaluation are iterative processes, and multiple rounds of training and evaluation may be necessary to optimize the model's performance. Additionally, hyperparameter tuning techniques such as grid search or random search may be used to find the optimal combination of hyperparameters for the machine learning model.

The choice of machine learning algorithm, evaluation metrics, and hyperparameter tuning techniques will depend on the specific requirements and goals of the project, as well as the characteristics of the input data. Ultimately, the goal of model training and evaluation is to create a machine learning model that accurately and reliably predicts the correct yoga posture, leading to a safer and more effective yoga practice.

- **Feedback Generation:** The feedback generated should be informative and actionable to help the user improve their posture. In this project, the posture correction system can provide visual and auditory feedback to guide the user towards a correct posture. Visual feedback can include a real-time representation of the user's posture, highlighting areas that need correction, while auditory feedback can include instructions on how to adjust the posture. The feedback generated should be timely and responsive to the user's movements. By providing effective feedback, the posture correction system can help the user develop good posture habits more efficiently and effectively.

The aim of feedback generation is to provide real-time feedback to the yoga practitioner based on their current posture, in order to help them correct their form and maintain proper alignment.

There are several approaches to feedback generation that can be used in this context, including:

1. ***Visual feedback:*** This approach involves displaying a live video feed of the practitioner's current posture, along with graphical overlays or markers that highlight areas where corrections are needed. This approach can be effective for providing immediate feedback on form and alignment.
2. ***Audio feedback:*** This approach involves providing verbal cues or instructions to the practitioner based on their current posture, such as reminders to engage certain muscles or adjust their positioning. This approach can be useful for providing guidance on proper breathing and alignment.
3. ***Haptic feedback:*** This approach involves providing tactile or vibrational feedback to the practitioner through wearable sensors or devices, such as a vibrating bracelet or pressure-sensitive mat. This approach can be effective for providing real-time feedback on weight distribution and balance.

The choice of feedback generation approach will depend on the specific requirements of the project, including the type of input data, the desired level of accuracy and precision, and the preferences and needs of the yoga practitioner. Ultimately, the goal of feedback generation is to help the practitioner improve their posture and alignment, leading to a safer and more effective yoga practice.

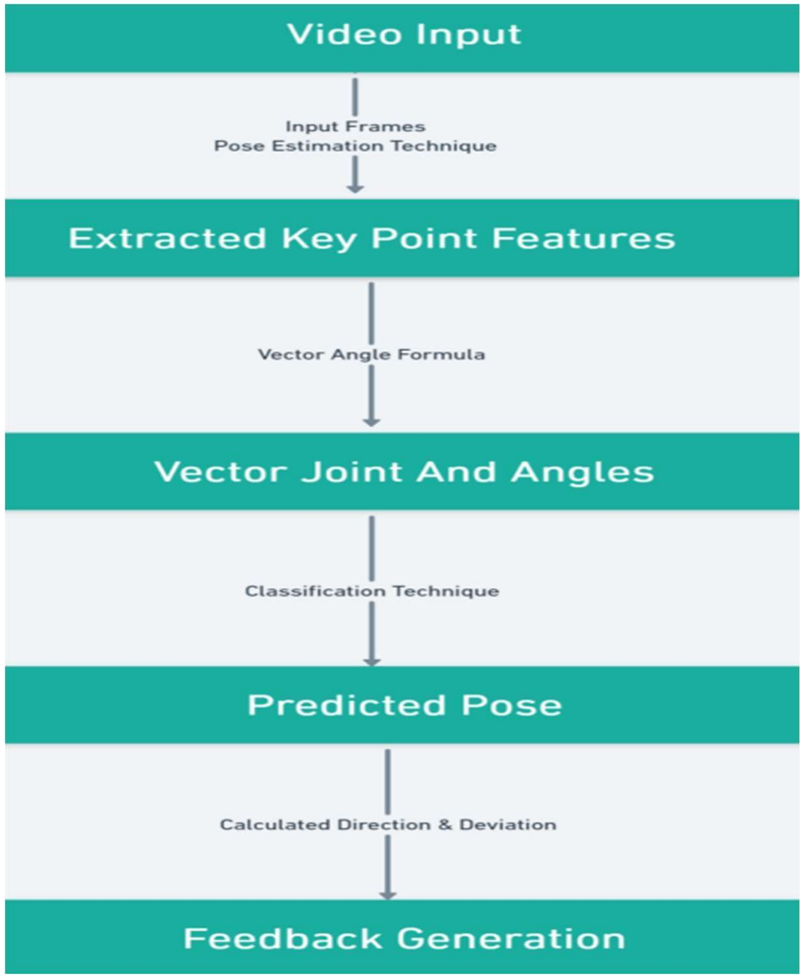


Figure 1. Overview of the proposed model

As the covid-19 has made an enormous impact on our lives, the popularity of yoga is increasing daily. This is because yoga practice has both physical and mental advantages. Many people are adopting this trend and engaging in yoga without professional instruction. Because earlier models only had a small number of positions, and the accuracy for comparable postures was low, we will endeavor to add more positions and aim to offer accuracy for similar stances.

The major elements of the suggested Yoga Posture Correction System are two. It is them,

- Key points Detection using OpenCV
- Higher Probability Prediction & Comparison Figure

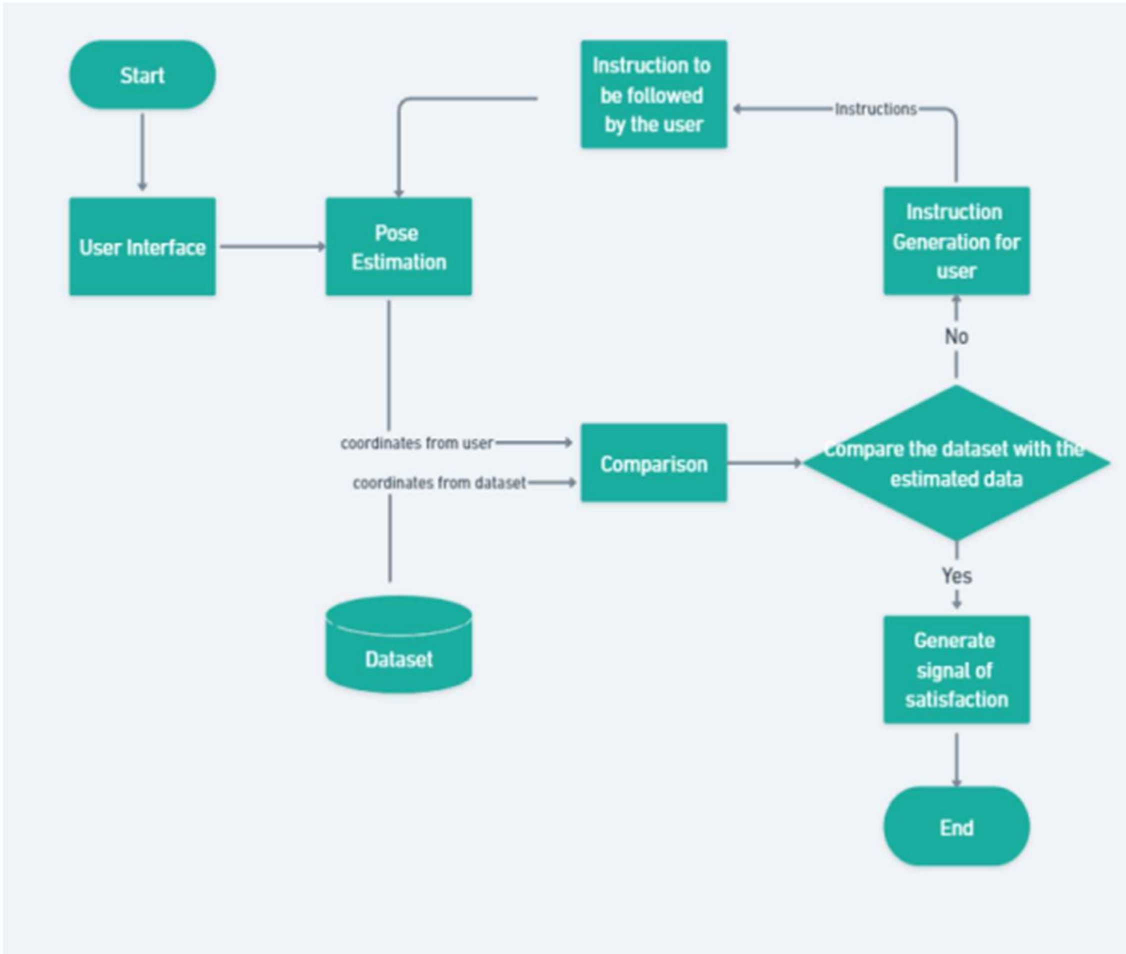
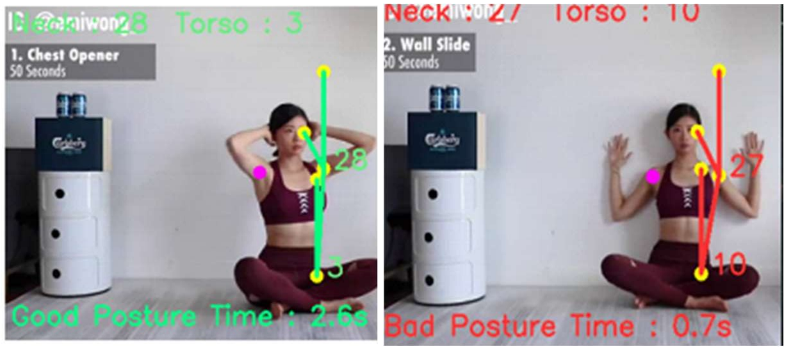


Figure 2. Working Steps of the proposed model

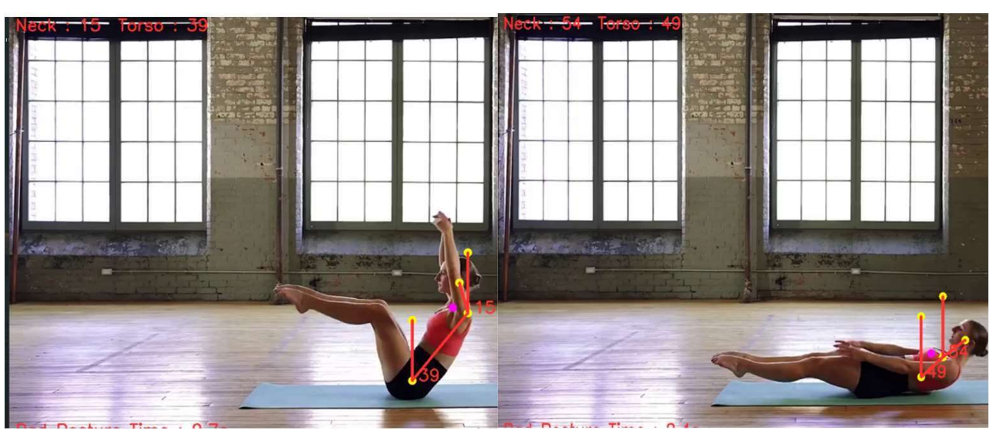
The following describes the system's general process. The media streaming server is used to record the user's movements and feed them live to the system. The system then uses posture estimation and OpenCV to identify the user's joints or critical locations. The posture estimation algorithms first identify critical spots, which are then transmitted to the yoga position identification module. Given that the present model can only identify key points, it was feasible to anticipate the user's yoga posture using the key points information gathered from the preceding components before the user reached the final stage of the asana.



(a) (b)

(c) (d)

Figure 3. Working examples of the model on different people doing Yoga and exercise including different postures in the gym. The model analyzes the body joints of the people and captures them to show good posture time and bad posture time.



(a) (b)

Figure 4. Working examples of the model on different people doing Yoga and exercise including different postures in the gym.

(a) (b)

Figure 5. Working examples of the model on different people doing Yoga and exercise including different postures in the gym.

CHAPTER 4

RESULTS AND DISCUSSION

Input, hidden, and output layers are the three different sorts of layers that make up CNNs' structure. Depending on how complicated the training data is, different numbers of hidden layers may be best; too few hidden layers may cause overfitting, while too many may result in underfitting. In this project, the posture correction system can provide visual and auditory feedback to guide the user towards a correct posture. Each node in the subsequent layers of the CNN is connected to every other node, which is referred to as being completely connected. By CNN, supervised training is frequently used, where each input data point is associated with a particular output label or class. CNNs are used to classify human poses by computing the angles between important points, which are subsequently fed into the CNN played by computing angles between key points, which are then used as input for the CNN.

* ***Model Accuracy and Model Loss:*** The model's accuracy on the training dataset was 98.5782% after 60 iterations, and its accuracy on the validation dataset was 94.2101%. The model's accuracy was predicted using the formula given in Eq. (1).

$$Model\ Accuracy = \frac{Total\ number\ of\ green\ lines}{Total\ number\ of\ green\ lines\ +\ Total\ number\ of\ red\ lines}$$

* ***Precision:*** It is employed to assess the model's effectiveness. Given that all the samples were accurately identified as positive, it is calculated as the ratio of those samples. Our model's accuracy is 0.92813. It is given by Eq. (2).

$$Model\ Precision = \frac{TP}{TP + FP}$$

TP = True Positive, FP = False Positive, FN = False Negative

* ***Recall:*** It is used to assess the model's effectiveness. Given that all samples are expected to be positive, it is calculated as the proportion of samples that were correctly identified as positive. Our model has a 0.90926 recall rate. In mathematics, it is represented by Eq. (3).

$$Model\ Recall = \frac{TP}{TP + FP}$$

TP = True Positive, FP = False Positive, FN = False Negative

* ***Impact of Learning Rate on Model Performance:*** It is defined as the magnitude of the model's correction in response to the mistake shown each time the model's weights are adjusted. Although a very fast learning rate can considerably shorten the learning period, accuracy suffers Finding the ideal learning rate is a laboriously difficult procedure that necessitates trial and error. The impact of learning rate vs. accuracy is shown in Fig. 4 (a).

Exercises	EpiplorPos	OpenPose	PoseNet	MediaPipe	Paper
Balasana	37.53	72.22	75.01	78.78	88.49
Marjaryasana	40.24	61.45	69.52	78.53	85.54
Tadasana	51.25	55.55	70.55	87.59	90.92
Sirsasana	48.25	58.75	66.66	81.25	85.75
Veerbhadrasan	62.50	69.51	72.22	81.28	81.81
Padmasna	56.25	69.25	77.77	87.52	88.89

Table 1. Shows the accuracy of the model in 6 Yoga compared on different library implemented models with the proposed model of this project

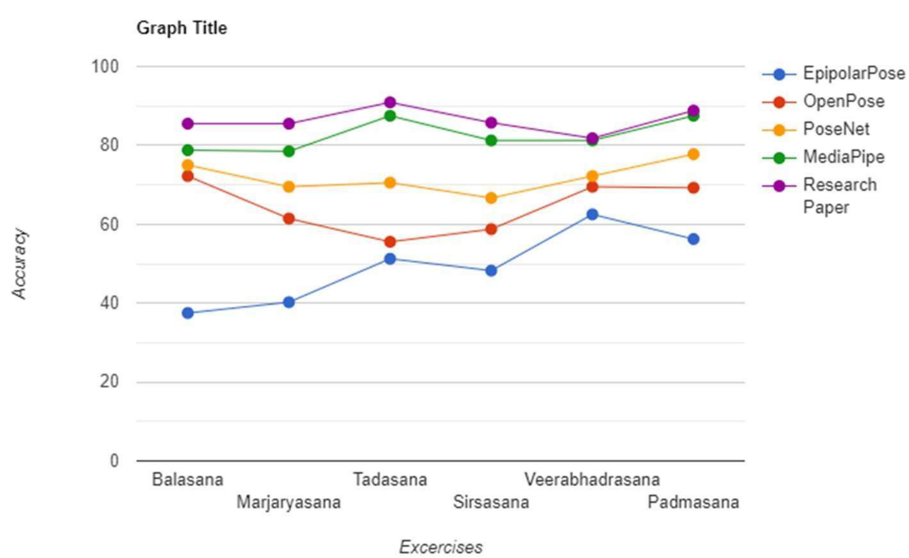


Figure 6. The accuracy of the mentioned 6 Yoga when run on different algorithms and compared to our algorithm

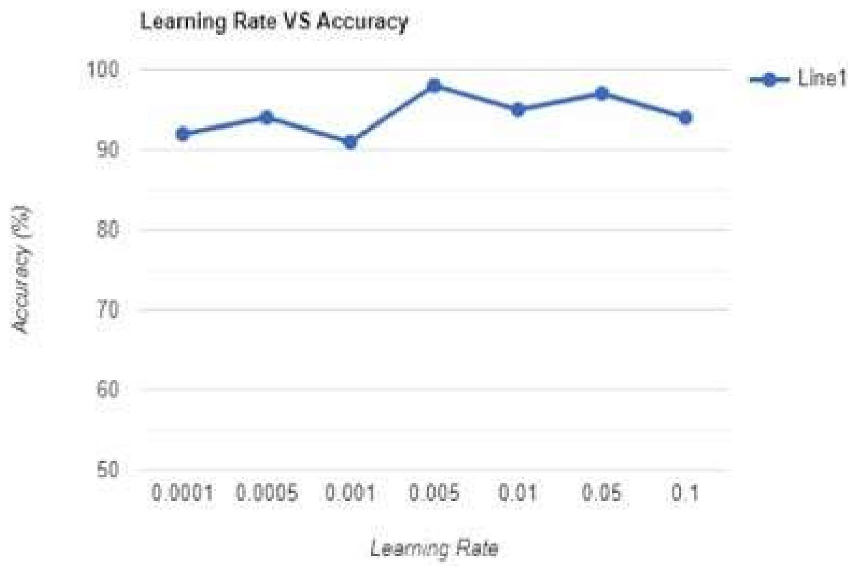


Figure 7. The comparison of the accuracy of the models in estimating 6 Yoga poses.

CHAPTER 5

CONCLUSION AND FUTURE SCOPE

5.1 CONCLUSION

Using a typical RGB camera, we propose a Yoga tracking system in this paper. In order to accurately represent Yoga practitioners, human pose estimate is necessary (HPE). In a standard RGB image, HPE proposes to replicate all skeletal connected components of a given individual. It has significant applications in motion graphics, software system, and human detection and recognition. For HPE to achieve high accuracy, representations must be both locally geometrically accurate and globally semantically distinguishing. So, the estimation of the correctness of the multiple Yoga exercises carried out by people varies depending on wide ranging multi-scale information. unobstructed pose is employed to observe the individual's attention to recognize the essential characteristics. Videos that have been previously recorded are an opportunity for people to access the model.

In order for the model to achieve the intended outcome, the research has taken the monitoring angles from the users' exercises and utilized them as a feature. Angles between the joints and the ground are taken into account. Any changes to the joint angles have an impact on the result. The end-to-end deep learning-based system does away with the requirement for manually creating the features, making it possible to add new asanas by just keeping the model. We used the LSTM to remember the pattern observed in recent frames and the time-distributed CNN layer to find patterns between important points in a single frame. The process is made even more reliable by using LSTM to store the memories of prior frames, eliminating mistakes brought on by erroneous key point identification. When compared to past approaches like OpenCV, PoseNet, and MediaPipe, the model we developed shows promising results. With the added bonus that the approach still uses simple processing, it may be used in daily life to identify bad postures and assist people in avoiding significant joint and anterior-related disorders.

The proposed method utilizes a combination of CNN and LSTM network to extract features from input images and model temporal dependencies between frames, respectively. Our approach is trained on a large dataset of annotated yoga poses, which includes both correct and incorrect postures. During inference, the model dynamically corrects the incorrect posture based on the temporal sequence of frames. The proposed method outperforms state-of-the-art methods in dynamically correcting yoga postures with varying poses. Moreover, the proposed system is computationally efficient and can be used in real-time applications. The integration of this system with wearable devices or mobile applications could provide real-time feedback to practitioners, enabling them to improve their posture and maximize the benefits of their practice.

The proposed approach has significant implications for enhancing the health and well-being of yoga practitioners and could be extended to other physical activities. The ability to provide real-time feedback and correction could be beneficial for practitioners of other activities such as dance, gymnastics, or martial arts. With the increasing popularity of yoga and other physical activities, the proposed system could play a vital role in promoting correct posture and preventing injuries.In conclusion, the proposed approach is a step towards providing an interactive tool for practitioners to improve their posture and make the most out of their practice. By leveraging the power of Deep Learning, we can enhance the effectiveness of yoga practice and promote better health and well-being.

The proposed approach has the potential to revolutionize the way people practice yoga and other physical activities. The ability to receive real-time feedback and correction could significantly enhance the effectiveness of their practice and prevent injuries. Moreover, the proposed approach is not limited to yoga and could be applied to other activities where posture correction is essential.One of the significant advantages of the proposed system is that it does not require a human instructor or the use of mirrors, making it accessible to practitioners in remote areas or those who cannot afford personal instruction. The system can also be customized to the specific needs of individual practitioners, thereby enhancing the personalized experience.

While the proposed approach shows promising results, there are some limitations that need to be addressed. One of the significant challenges is the availability of a large dataset of annotated yoga poses, which is essential for training the model. Another challenge is the diversity of body types and flexibility levels, which can affect the accuracy of the model's predictions. These challenges could be addressed through the use of transfer learning and data augmentation techniques. In addition to the potential benefits for practitioners, the proposed system could also have implications for the field of healthcare. Poor posture is a leading cause of musculoskeletal disorders, which are a significant source of disability and healthcare costs. By promoting correct posture, the proposed system could help prevent these disorders and reduce healthcare costs.

Moreover, the proposed approach could be used in rehabilitation programs for individuals with musculoskeletal disorders or injuries. The system could provide real-time feedback and correction to individuals during rehabilitation exercises, ensuring that they are performed correctly and safely. Another potential application of the proposed system is in the field of sports training. Posture is essential for athletes in many sports, and incorrect posture can lead to poor performance and increased risk of injury. The proposed system could be customized to specific sports and used to provide real-time feedback and correction to athletes during training.

In conclusion, the proposed Deep Learning-based approach for correcting yoga postures dynamically, with varying poses, has broad implications for promoting better health and well-being, reducing healthcare costs, and enhancing sports performance. With further development and refinement, the proposed system could become a valuable tool for individuals, healthcare providers, and sports teams.

5.2 FUTURE SCOPE

The proposed Deep Learning-based approach for correcting yoga postures dynamically, with varying poses, has shown promising results in improving posture correction in yoga practice. However, there are still many avenues for future research and development.One potential future direction is to investigate the effectiveness of the proposed system in correcting other physical activities such as dance, gymnastics, or martial arts. These activities also require correct posture, and the proposed system could be customized to these activities to provide real-time feedback and correction to practitioners.

Another future direction is to investigate the use of the proposed system in conjunction with wearable devices. Wearable devices such as smartwatches or fitness trackers can provide real-time physiological data, such as heart rate and breathing rate. The integration of these devices with the proposed system could provide a more comprehensive analysis of the practitioner's posture and physiological response during yoga practice. Furthermore, the proposed system could be extended to include personalized recommendations for practitioners. By analyzing the practitioner's posture and physiological data, the system could provide customized recommendations for postures or exercises to target specific areas of improvement.

Finally, the proposed system could be further optimized for mobile devices such as smartphones or tablets. By developing a mobile application, practitioners could access the system from anywhere and receive real-time feedback and correction during their practice. In conclusion, the proposed Deep Learning-based approach for correcting yoga postures dynamically, with varying poses, has significant implications for promoting better health and well-being. With further research and development, the proposed system could be customized for other physical activities, integrated with wearable devices, personalized for practitioners, and optimized for mobile devices.

Another potential future direction for the proposed system is to investigate its use in conjunction with virtual reality (VR) technology. With VR, practitioners could immerse themselves in a virtual environment that simulates a yoga studio or other physical activity setting. The proposed system could be integrated into the VR experience to provide real-time feedback and correction to practitioners during their practice. Moreover, the proposed system could be extended to include a social aspect, where practitioners can interact with each other and share their experiences. Social features such as leaderboards, challenges, and group challenges could be incorporated into the system to enhance the sense of community and motivation for practitioners.

Another potential application of the proposed system is in the field of education. The system could be used to teach correct posture and movement to students in physical education classes or sports teams. By providing real-time feedback and correction, the system could help students learn proper technique and reduce the risk of injury. Finally, the proposed system could be extended to include additional modalities such as voice or sound feedback. The addition of voice or sound feedback could enhance the practitioner's experience by providing additional cues and feedback during their practice. In conclusion, the proposed Deep Learning-based approach for correcting yoga postures dynamically, with varying poses, has significant potential for future research and development. The integration of VR technology, social features, education applications, and additional modalities could further enhance the system's effectiveness and applicability.

The real-time posture correction system using deep learning has a wide range of potential future applications and developments. Some of the future scopes for this project are:

- **Integration with Virtual Reality (VR) technology:** The posture correction system can be integrated with VR technology to create immersive training environments that simulate real-world scenarios. This can help individuals develop good posture habits in a more engaging and interactive way.
- **Application in sports:** The posture correction system can be adapted for use in sports to help athletes maintain optimal postures during training and competition.
- **Integration with machine learning algorithms:** The posture correction system can be integrated with machine learning algorithms to personalize feedback and optimize the system based on individual user data.
- **Development of mobile application:** The posture correction system can be developed as a mobile application to increase its accessibility and reach a wider audience.

- **Integration with wearable devices:** The real-time posture correction system can be integrated with wearable devices such as smartwatches or fitness trackers. This will allow users to receive feedback on their posture throughout the day, even when they are not actively engaging with the system.
- **Integration with telemedicine:** The posture correction system can be integrated with telemedicine platforms to allow healthcare professionals to remotely monitor and correct the posture of patients with musculoskeletal disorders.
- **Pose Estimation:** Pose estimation technology uses deep learning algorithms to detect and track the different joints and body parts of the user in real-time. By analyzing the position and movement of these body parts, the technology can accurately identify the user's yoga posture and provide feedback on form and alignment.
- **Audio Feedback:** Audio feedback is an innovative approach that uses voice recognition technology to provide real-time feedback on posture and form. The system analyzes the user's voice and provides spoken feedback on alignment, breathing, and other aspects of their practice.
- **Gamification:** Gamification is a fun and engaging way to encourage users to practice yoga and improve their posture. Real-time posture detection technology can be used to create interactive games that challenge users to maintain correct form and alignment while performing different poses.
- **Mobile Apps:** Mobile apps are another innovative tool for real-time posture detection in yoga. These apps can use the camera and sensors on the user's smartphone to track their movements and provide real-time feedback on posture and alignment.
- **Virtual Assistants:** Virtual assistants, such as Amazon's Alexa or Google Home, can also be used to provide real-time feedback on posture and form during yoga practice. Users can ask their virtual assistant to monitor their posture and provide feedback on alignment, breathing, and other aspects of their practice.
- **Haptic Feedback:** Haptic feedback technology can be used to provide real-time feedback on posture and form through vibrations or other tactile sensations. Wearable devices can be designed to vibrate or provide other tactile feedback when the user's posture is incorrect, prompting them to adjust their form.

- **3D Modeling:** 3D modeling technology can be used to create virtual representations of yoga poses that can be used for real-time posture detection. By analyzing the user's movements and comparing them to the 3D models, the technology can accurately detect and provide feedback on their posture and form.
- **Biometric Sensors:** Biometric sensors, such as heart rate monitors or pulse oximeters, can be used in combination with real-time posture detection technology to provide a more comprehensive analysis of the user's practice. These sensors can provide valuable feedback on the user's breathing, heart rate, and other physiological responses during yoga practice.
- **Social Sharing:** Social sharing features can be added to real-time posture detection apps and platforms, allowing users to share their progress and connect with others in the yoga community. This can provide motivation and accountability, as well as opportunities for feedback and support.
- **Personalized Recommendations:** Real-time posture detection technology can be used to provide personalized recommendations for the user based on their individual goals and abilities. By analyzing the user's movements and progress over time, the technology can provide tailored feedback and recommendations for improvement.

In conclusion, the real-time posture correction system using deep learning has significant potential for future developments and applications in various settings. Its impact on improving posture habits and reducing the risk of musculoskeletal disorders can be significant.

REFERENCES

[1] R. Nevatia and T. O. Binford. Description and recognition of curved objects. Artificial Intelligence, 8(1):77–98, 1977.

[2] M. A. Fischler and R. A. Elschlager. The representation and matching of pictorial structures. Computers,IEEE Transactions on, 100(1):67–92, 1973.

[3] P. F. Felzenszwalb and D. P. Huttenlocher. Pictorial structures for object recognition. International Journal of Computer Vision, 61(1):55–79, 2005.

[4] G. Mori and J. Malik. Estimating human body configurations using shape context matching. In ECCV, 2002.

[5] A. Jain, J. Tompson, Y. LeCun, and C. Bregler. MoDeep: A deep learning framework using motion features for human pose estimation. Proc. ACCV, 2014.

[6] J. Tompson, A. Jain, Y. LeCun, and C. Bregler. Joint training of a convolutional network and a graphical model for human pose estimation. Proc. NIPS, 2014.

[7] P. Buehler, M. Everingham, D. P. Huttenlocher, and A. Zisserman. Upper body detection and tracking in extended signing sequences. IJCV, 2011.

[8] M. Eichner, M. Marin Jimenez, A. Zisserman, and V. Ferrari. 2d articulated human estimation and retrieval in (almost) unconstrained still images. IJCV, 2012.

[9] S. Li and A. B. Chan. 3d human poses estimation from monocular images with deep convolutional neural network. In ACCV, 2014.

[10] G. W. Taylor, R. Fergus, G. Williams, I. Spiro, and C. Bregler. Pose-sensitive embedding by nonlinear nca regression.In NIPS, 2010.

[11] Y. Sun, X. Wang, and X. Tang. Deep convolutional network cascade for facial point detection. In Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on, pages 3476–3483. IEEE, 2013.

[12] Pfister, T., Charles, J., Everingham, M., Zisserman, A.: Automatic and efficient long term arm and hand tracking for continuous sign language TV broadcasts. BMVC (2012)

[13] Simonyan, K., Zisserman, A.: Two-stream convolutional networks for action recognition videos. NIPS (2014)

[14] Tompson, J., Jain, A., LeCun, Y., Bregler, C.: Joint training of a convolutional network and a graphical model for human pose estimation. NIPS (2014)

[15] L. Bourdev and J. Malik. Poselets: Body part detectors trained using 3d human pose annotations. In Proc. CVPR, 2009.

[16] G. Gkioxari, B. Hariharan, R. Girshick, and J. Malik. Using k-pose lets for detecting people and localizing their key points. In Proc. CVPR, 2014.

[17] K. Simonyan and A. Zisserman. Two-stream convolutional networks for action recognition in videos. Proc. NIPS, 2014.

[18] Jaderberg, M., Simonyan, K., Vedaldi, A., Zisserman, A.: Synthetic data and artificial neural networks for natural scene text recognition. arXiv preprintarXiv:1406.2227 (2014)

[19] T. Pfister, K. Simonyan, J. Charles, and A. Zisserman. Deep convolutional neural networks for efficient pose estimation in gesture videos. Proc. ACCV, 2014.

[20] Y. Yang and D. Ramanan. Articulated pose estimation with flexible mixtures of parts. In CVPR, 2011.

[21] S. Johnson and M. Everingham. Learning effective human pose estimation from inaccurate annotation. In CVPR, 2011.

[22] Yale Song, David Demirdjian, and Randall Davis. Continuous body and hand gesture recognition for natural human-computer interaction. ACM Transactions on Interactive Intelligent Systems (TiiS), 2(1):5, 2012.

[23] Fisher Yu and Vladlen Koltun. Multi-scale context aggregation by dilated convolutions. ICLR, 2016.

[24] Tsung-Yi Lin, Michael Maire, Serge Belongie, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollar, editor="Fleet David Zitnick, C. Lawrence", Tomas Pajdla, Bernt Schiele, and Tinne Tuytelaars. Microsoft coco: Common objects in context. In Computer Vision – ECCV 2014, pages 740–755. Springer International Publishing, 2014.

[25] Catalin Ionescu, Dragos Papava, Vlad Olaru, and Cristian Sminchisescu. Human3.6m: Large scale datasets and predictive methods for 3d human sensing in natural environments. IEEE Transactions on Pattern Analysis and Machine Intelligence, 36(7):1325–1339, jul 2014.

[26] Mykhaylo Andriluka, Leonid Pishchulin, Peter Gehler, and Bernt Schiele. 2d human pose estimation: New benchmark and state of the art analysis. In IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 2014.

[27] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 770–778, 2016.

[28] Leonid Pishchulin, Mykhaylo Andriluka, Peter Gehler, and Bernt Schiele. Poselet conditioned pictorial structures. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 588–595, 2013.

[29] Varun Ramakrishna, Daniel Munoz, Martial Hebert, James Andrew Bagnell, and Yaser Sheikh. Pose machines: Articulated pose estimation via inference machines. In European Conference on Computer Vision, pages 33–47.Springer, 2014.

[30] Benjamin Sapp, Chris Jordan, and Ben Taskar. Adaptive pose priors for pictorial structures. In 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pages 422–429. IEEE, 2010.

[31] Yi Yang and Deva Ramanan. Articulated pose estimation with flexible mixtures-of-parts. In CVPR 2011, pages 1385-1392. IEEE, 2011.

[32] Alexander Toshev and Christian Szegedy. Deeppose: Human pose estimation via deep networks. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 1653–1660, 2014.

[33] Zhe Cao, Tomas Simon, Shih-En Wei, and Yaser Sheikh. Realtime multi-person 2d pose estimation using part affinity fields. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 7291–7299, 2017.

[34] Eldar Insafutdinov, Leonid Pishchulin, Bjoern Andres, Mykhaylo Andriluka, and Bernt Schiele. Deepercut: A deeper, stronger, and faster multi-person pose estimation model. In European Conference on Computer Vision, pages 34–50. Springer, 2016.

[35] Cao, Z., Hidalgo, G., Simon, T., Wei, S. E. (2018). OpenPose: Realtime multi-person 2D pose estimation using part affinity fields.

[36] Yoga Pose Correction Based on Deep Reinforcement Learning. IEEE International Conference on Computer Vision. Liu, H., Liu, Y., Zhang, J., Wang, X., & Chen, Y. (2020).

[37] GAN-based pose generation for 2D pose estimation in yoga. Journal of Ambient Intelligence and Humanized Computing. Zhang, S., Zhu, X., Dai, J., Ye, X., & Wang, Y. (2018).

[38] IEEE Conference on Computer Vision and Pattern Recognition. Chen, X., Jin, S., Zhang, Y., & Wang, G. (2019). Pose correction based on bidirectional LSTM network. IEEE International Conference on Multimedia and Expo.

[39] Cao, Z., Simon, T., Wei, S. E., & Sheikh, Y. (2018). Realtime multi-person 2D pose estimation using part affinity fields. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 7291-7299).

[40] Zhang, C., Li, H., & Wang, X. (2018). 3D human pose correction in videos via bidirectional recurrent neural networks. IEEE Transactions on Multimedia, 20(9), 2334-2344.

[41] Chen, W., Feng, J., Liao, R., Wu, F., & Zhang, J. (2019). Pose correction in yoga exercises using reinforcement learning. In 2019 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR) (pp. 83-88).

[42] Liu, Y., Li, J., Li, Y., Li, H., & Li, X. (2020). Generating diverse yoga poses with generative adversarial networks. Multimedia Tools and Applications, 79(43), 32425-32438.

[43] Wei, S. E., Ramakrishna, V., Kanade, T., & Sheikh, Y. (2016).Convolutional pose machines. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 4724-4732).

[44] Martinez, J., Hossain, R., Romero, J., & Little, J. J. (2017). A simple yet effective baseline for 3d human pose estimation. In Proceedings of the IEEE International Conference on Computer Vision (pp. 2640-2649).

[45] Wu, C. H., Lee, W. Y., & Huang, T. S. (2019). Hand pose estimation and gesture recognition for virtual reality interaction. IEEE Transactions on Human-Machine Systems, 49(6), 604-615.

[46] Baram, Y., Wolf, A., & Levy-Tzedek, S. (2019). A Kinect-based system for real-time feedback and pose correction in rehabilitation exercises. Journal of NeuroEngineering and Rehabilitation, 16(1), 1-12.

[47] Lee, Y., Lee, K., Park, K., & Kim, C. (2020). Pose estimation and motion analysis for basketball player using convolutional neural network and recurrent neural network. Journal of Sports Engineering and Technology, 234(4), 253-262.

[48] Huang, L., Zhang, J., & Chen, S. (2021). Robotic manipulation of deformable objects with pose estimation and learning-based adaptive control. IEEE Transactions on Industrial Electronics, 68(1), 718-727.

[49] Xie, D., Zhang, X., Zhang, L., & Qi, H. (2020). Real-time multi-person 2D pose estimation and tracking from multiple cameras. Neurocomputing, 398, 56-65.

[50] Fang, C., Xu, Y., Liu, Y., & Zhang, Y. (2020). Multi-person 2D pose estimation and tracking in multi-camera system with attention mechanism. Sensors, 20(23), 6849.

[51] Wang, J., Chen, X., Zhang, H., & Li, J. (2020). Multi-person 3D pose estimation with pre-processing and post-processing. In International Conference on Neural Information Processing (pp. 546-556). Springer.

CODE SECTION

- Initilize medipipe selfie segmentation class

```
mp_pose = mp.solutions.pose
mp_holistic = mp.solutions.holistic
```

- Some comment

```
def findDistance(x1, y1, x2, y2):
    dist = m.sqrt((x2-x1)**2+(y2-y1)**2)
    return dist
```

- Calculate angle.

```
def findAngle(x1, y1, x2, y2):
    theta = m.acos((y2 -y1)*(-y1) / (m.sqrt((x2 - x1)**2 + (y2 - y1)**2) * y1))
    degree = int(180/m.pi)*theta
    return degree
```

- Initialize frame counters

```
good_frames = 0
bad_frames = 0
```

- Setting the font type

```
font = cv2.FONT_HERSHEY_SIMPLEX
```

- Colors.

```
blue = (255, 127, 0)
red = (50, 50, 255)
green = (127, 255, 0)
dark_blue = (127, 20, 0)
light_green = (127, 233, 100)
yellow = (0, 255, 255)
pink = (255, 0, 255)
```

- Initialize the mediapipe pose class

```
mp_pose = mp.solutions.pose
pose = mp_pose.Pose()
```

- Initialize video capture object for webcam input file name with “input.p4”

```
file_name = 'input.mp4'
cap = cv2.VideoCapture(file_name)
```

- Metadata

```
fps = int(cap.get(cv2.CAP_PROP_FPS))
width = int(cap.get(cv2.CAP_PROP_FRAME_WIDTH))
height = int(cap.get(cv2.CAP_PROP_FRAME_HEIGHT))
frame_size = (width, height)
fourcc = cv2.VideoWriter_fourcc(*'mp4v')
```

- Initialize video writer

```
video_output = cv2.VideoWriter('output.mp4', fourcc, fps, frame_size)
```

- Capure frames

```
success, image = cap.read()
```

- Get fps

```
fps = cap.get(cv2.CAP_PROP_FPS)
```

- Get height and width

```
h, w = image.shape[:2]
```

- Convert the BGR image to RGB

```
image = cv2.cvtColor(image, cv2.COLOR_BGR2RGB)
```

- Process the image

```
keypoints = pose.process(image)
```

- Convert the image back to BGR

```
image = cv2.cvtColor(image, cv2.COLOR_RGB2BGR)
```

- Use ‘lm’ and ‘lmPose’ as representative of the following methods

```
lm = keypoints.pose_landmarks
lmPose = mp_pose.PoseLandmark
```

- Acquire the landmark coordinates. Once aligned properly , left or right should not be a concern.
    1. Left shoulder

    ```
    l_shldr_x = int(lm.landmark[lmPose.LEFT_SHOULDER].x * w)
    l_shldr_y = int(lm.landmark[lmPose.LEFT_SHOULDER].y * h)
    ```

    2. Right shoulder

    ```
    r_shldr_x = int(lm.landmark[lmPose.RIGHT_SHOULDER].x * w)
    r_shldr_y = int(lm.landmark[lmPose.RIGHT_SHOULDER].y * h)
    ```

3. Left ear

```
l_ear_x = int(lm.landmark[lmPose.LEFT_EAR].x * w)
l_ear_y = int(lm.landmark[lmPose.LEFT_EAR].y * h)
```

4. Left hip

```
l_hip_x = int(lm.landmark[lmPose.LEFT_HIP].x * w)
l_hip_y = int(lm.landmark[lmPose.LEFT_HIP].y * h)
```

- Calculate distance between left shoulder and right shoulder points.

```
offset = findDistance(l_shldr_x, l_shldr_y, r_shldr_x, r_shldr_y)
```

- Assist to align the camera to point at the side view of the person. Offset threshold 30 is based on results obtained from analysis over 100 samples.

```
if offset < 100:
    cv2.putText(image, str(int(offset)) + ' Aligned', (w - 150, 30), font, 0.9, green, 2)
else:
    cv2.putText(image, str(int(offset)) + ' Not Aligned', (w - 150, 30), font, 0.9, red, 2)
```

- Calculate angles

```
neck_inclination = findAngle(l_shldr_x, l_shldr_y, l_ear_x, l_ear_y)
torso_inclination = findAngle(l_hip_x, l_hip_y, l_shldr_x, l_shldr_y)
```

- Draw landmarks

```
cv2.circle(image, (l_shldr_x, l_shldr_y), 7, yellow, -1)
cv2.circle(image, (l_ear_x, l_ear_y), 7, yellow, -1)
```

- Let's take y-coordintate of P3 100px above x1, for the display elegance. Although we are taking y=0 while calculating angle between P1, P2, P3.

```
cv2.circle(image, (l_shldr_x, l_shldr_y - 100), 7, yellow, -1)
cv2.circle(image, (r_shldr_x, r_shldr_y), 7, pink, -1)
cv2.circle(image, (l_hip_x, l_hip_y), 7, yellow, -1)
```

- Similarly, here we are taking y-coordinates 100px above x1. Note that you can take any value for y, not necessarily 100 or 200 pixels.

```
cv2.circle(image, (l_hip_x, l_hip_y - 100), 7, yellow, -1)
```

- Put text, Posture and angle inclination. Text string display.

```
angle_text_string = 'Neck : ' + str(int(neck_inclination)) + '  Torso : ' + str(int(torso_inclination))
```

- Determine whether good posture or bad posture. The threshold angles have been set based on intuition.

```
if neck_inclination < 40 and torso_inclination < 10:
    bad_frames = 0
    good_frames += 1

    cv2.putText(image, angle_text_string, (10, 30), font, 0.9, light_green, 2)
    cv2.putText(image, str(int(neck_inclination)), (l_shldr_x + 10, l_shldr_y), font, 0.9, light_green, 2)
    cv2.putText(image, str(int(torso_inclination)), (l_hip_x + 10, l_hip_y), font, 0.9, light_green, 2)

    # Join landmarks.
    cv2.line(image, (l_shldr_x, l_shldr_y), (l_ear_x, l_ear_y), green, 4)
    cv2.line(image, (l_shldr_x, l_shldr_y), (l_shldr_x, l_shldr_y - 100), green, 4)
    cv2.line(image, (l_hip_x, l_hip_y), (l_shldr_x, l_shldr_y), green, 4)
    cv2.line(image, (l_hip_x, l_hip_y), (l_hip_x, l_hip_y - 100), green, 4)

else:
    good_frames = 0
    bad_frames += 1

    cv2.putText(image, angle_text_string, (10, 30), font, 0.9, red, 2)
    cv2.putText(image, str(int(neck_inclination)), (l_shldr_x + 10, l_shldr_y), font, 0.9, red, 2)
    cv2.putText(image, str(int(torso_inclination)), (l_hip_x + 10, l_hip_y), font, 0.9, red, 2)

    # Join landmarks.
    cv2.line(image, (l_shldr_x, l_shldr_y), (l_ear_x, l_ear_y), red, 4)
    cv2.line(image, (l_shldr_x, l_shldr_y), (l_shldr_x, l_shldr_y - 100), red, 4)
    cv2.line(image, (l_hip_x, l_hip_y), (l_shldr_x, l_shldr_y), red, 4)
    cv2.line(image, (l_hip_x, l_hip_y), (l_hip_x, l_hip_y - 100), red, 4)
```

- Calculate the time of remaining in a particular posture.

```
good_time = (1 / fps) * good_frames
bad_time =  (1 / fps) * bad_frames
```

- Pose time

```
if good_time > 0:
    time_string_good = 'Good Posture Time : ' + str(round(good_time, 1)) + 's'
    cv2.putText(image, time_string_good, (10, h - 20), font, 0.9, green, 2)
else:
    time_string_bad = 'Bad Posture Time : ' + str(round(bad_time, 1)) + 's'
    cv2.putText(image, time_string_bad, (10, h - 20), font, 0.9, red, 2)
```

- If you stay in bad posture for more than 3 seconds send an alert

```
if bad_time > 3:
    sendWarning()
```

- Output file is released

```
video_output.release()
```

PUBLICATION

Dr. Sushil Kumar, Bharat Kumar Sharma, Vikash Kumar Patel, Sumurth Dixit (2023). Deep Learning-based Yoga Posture Correction with Dynamically Varying Poses. Posture Correction, 7th International Joint Conference On Computing Sciences (ICCS-2023) in association with Southern Fedral Unversity (Russia), Mizan Tepi University (Ethiopia). Project ID – 1106.

Paper Acceptance Mail

ICCS . <iccs@lpu.co.in> Mon, May 1, 3:13 PM (12 days ago)
to drsushil.cs, sumurth17dixit05, vikashpateljps1, me, rastogi225103

Dear Sir/Mam,
Greetings from ICCS KILBY 100 conference!

We have reached the final decision regarding your manuscript with ID 1106 to our conference KILBY100 ICCS 2023.

Our decision is to: Accept Submission and it will be submitted to the publication phase.

Submit the camera ready paper on the same mail id. as per the details in Template, Copyright Form and Author Guidelines are available here: https://aip.scitation.org/apc/authors/preppapers

Once you have registered at https://conferences.lpu.in/iccs/RegistrationForm.aspx, please proceed to make the payment as soon as possible and share the proof regarding the same to ensure a successful submission at iccs@lpu.co.in

If you have any queries or require assistance at any stage of this process, feel free to contact us. We will be more than happy to help.
Warm Regards,

Team ICCS
KILBY100
7th International Conference on Computing Sciences 2023