National College of Ireland

Yoga Pose Prediction Using InceptionV3 an Transfer Learning Approach

MSc Research Project
MSc Data Analytics

Monika Maheswari Baskar
Student ID: x22200169

School of Computing
National College of Ireland

Supervisor: Vikas Tomer

National College of Ireland

MSc Project Submission Sheet

School of Computing

Student Name:	Monika Maheswari Baskar		
Student ID:	X22200169		
Programme:	MSc Data Analytics	**Year:**	2023-2024
Module:	MSc Research Project		
Supervisor:	Vikas Tomer		
Submission Due Date:	16/09/2024		
Project Title:	Yoga Pose Prediction Using InceptionV3 an Transfer Learning Approach		
Word Count:	6259	**Page Count**	20

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.
ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature: Monika Maheswari Baskar

Date: 16/09/2024

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST

Attach a completed copy of this sheet to each project (including multiple copies)	✓
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	✓
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	✓

Assignments that are submitted to the Programme Coordinator Office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Yoga Pose Prediction Using InceptionV3 an Transfer Learning Approach

Monika Maheswari Baskar
X22200169

Abstract

This study recognizes human kinematics most accurately to help people have a good and healthy life in future by recognising the human posture using the InceptionV3 algorithm and Transfer learning. The field of computer vision has a difficult problem while estimating human posture error in posture can lead to discomfort and immediate issues. Identifying a person's posture in the photograph might be a little bit difficult due to the image size, resolution, lighting, background clutter, apparel and environment. The prediction is based on human skeleton motion through joint position using Mediapipe which detects the key points of body, foot, hand and face from single image. In order to predict whether the yoga postures are correct or wrong with the input image, the input image is changed to skeleton to get the performance metrics accurately. This research is fully based on predicting the yoga posture whether it is correct or incorrect by maximizing angles of 38 different yoga asanas through accurate prediction. This study in cooperates the CROSS-INDUSTRY STANDARD PROCESS FOR DATA MINING Methodology. The dataset was taken from Kaggle consists of 38 yoga asanas after doing the pre-processing After the pre- processing steps the images are augmented and converted into skeleton keyframes using the OpenCV and Mediapipe. The model used to predict the data is the InceptionV3 model which is a transfer learning approach and jupyter notebook is used for implementation. The models performance was evaluated using metrics such as accuracy, precision, recall and F1-score the training and validation accuracy was achieved as 96.03% and it also compared the STF-ResNet, AdaBoost, InceptionV3 and VGG model with accuracy around 77% to 93% respectively. Also, to provide the evaluation more compressively confusion matrix and classification report is provided. In recent years, yoga has been practiced by many people in terms of relaxation and for a peaceful mind but the problem with yoga is that like any other activity it is critical to do it because poor posture can make a yoga session unproductive and even detrimental. The proposed system aims to provide a comprehensive solution for individuals looking to enhance their yoga practice and improve their overall health and wellbeing.

1.Introduction

1.1 Research Background:

Exercise and physical activity can have both short term and long-term benefits on physical health. In such case posture is one of the most important aspect of exercise to get the benefits of those postures if there is any error in posture leads to sprains and strains in ligaments, muscles are caused also by stretching the body improperly or performing inappropriate posture can be too harmful for once life which also leads to immediate discomfort and long-term chronic challenges.

1.2 Research Question:

Measuring the impact of utilizing transfer learning and InceptionV3 model to predict the accurate yoga posture in order to develop the benefits of yoga.

1.3 Problem Solution:

To address such a problem, coordinate points were obtained from the human pose using the media pipe which helps to detect the human body, foot, hand and facial key points in the single picture and based on the key joint points it draws the skeleton as well. The CNN an InceptionV3 algorithm is used to classify this scenario where the angles have been measured and posture have been adjusted to improve the correct yoga posture. The purpose of this project is to predict whether the yoga posture is correct or wrong by analysing the angles of each and every yoga posture and maximizing the benefits of yoga by providing the most accurate prediction.

1.4 Novelty of the Study:

This study has the novelty by implementing the InceptionV3 algorithm, an Transfer learning approach with 38 different yoga asanas where all the asanas have been predicted correctly even by giving the normal asana image as the skeleton image with the respective names.

1.5 Document Structure:

- **Abstract**: Detailly explains what this project is all about.
- **Introduction:** Explains about what is the research question, novelty of the study, the background of this project and a solution.
- **Literature Review:** Provides the context of the current state of benefits of yoga, what are the algorithms they have used before and using InceptionV3 how the different study has been implemented.
- **Methodology:** Explains detailly about the project and what the methods are done step by step.
- **Design & Specification:** Explain the flow of this study and what models are used in detail.
- **Implementation:** Explains what has been done in the code with respective code snippets and necessary visualizations.
- **Evaluation & Results:** Based on the project what is the total accuracy, precision, recall, f1 score.
- **Conclusion & Future Work:** Why took this project and what can be implemented by using this InceptionV3 algorithm in future.

2.Related Work

2.1 Benefits of doing Yoga:

In order to improve physical, mental and social health yoga is practiced significantly to get long term benefits. This study examines Akdeniz et al. (2023) by doing yoga it helps people to reduce pain, enhance flexibility, improve sleep quality and a better stress management. The data is collected through semi-structured interviews with 18 participants in which 2 are men

and the rest are women's who practiced yoga regularly with the cause of physical, mental and social benefits. The drawbacks of this study is it only covers the response based on the 18 participants which could not be generalized. Also, the results are limited to certain institutions. The future work of this study will be exploring different yoga effects. In the similar way Mahindru et al.(2023) examines both the mental and physical health by doing yoga regularly it reduces the symptoms of depression and anxiety, improves sleep quality and decreases healthcare costs. This study is all about how to improve the physical and mental health by using some of the HPA axis functioning which helps to increase in endorphin production and yoga is used for various mental health conditions such as schizophrenia, disorders and sleep disturbances. Regarding the physical and mental health there are some search strategy databases namely PubMed, Google scholar and Medline and they did a research based on recent publications. The drawback of this paper is the lack of understanding of the yoga types and amount of yoga needed for mental health benefits. The future work of this project is enhancing training for healthcare professionals to implement a specific and the structured yoga therapies in the clinical world. Based on the previous paper limitations Patel et al.(2023) have examined the lack of yoga types and tells how much yoga is necessary in order to get the benefits of mental health. The novelty of this paper is promoting the mental and physical health during the covid-19 pandemic by using the Hatha Yoga which helps to gives a brief about our pulmonary health, boosting health and reducing stress the main reason for this research is due to their high risk f virus transmission as a result many turned to do yoga in their comfort of their home to maintain a good health.This study claims three main information about the breathing exercise, sustained yoga practice and reducing stress to contribute mental health. The limitations and future work of this study includes that they are facing some challenges to generalize the yoga practices based on individuals and should explore the virtual yoga sessions. This study examines the equavation between the yoga practice and lifestyle choices specified to India also it tells us what is choices of food , physical and leisure activity among yoga influences. Patel et al.(2023) took this study because of the impact of yoga on lifestyle choices specifically for the Indian people. Therefore this study targets 551 yoga practitioners who do yoga both online and offline methods and taking the sample from them. Including the yoga practice they have taken samples with food consumption and physical activities too. Based on the previous research the generalized feedback is limited in this research. The limitation insists that the self paced data introduces a response bias and the future work includes exploring different styles of the yoga practice in various groups. This paper is all about investigating the 4 week yoga class for cancer survivors based on the cancer related fatigue(CRF) by implementing some physical activity such as yoga, walking and more. Lin et al.(2023) took this study because CRF is a common issue among cancer people which affects their daily quality of life. All the participants were assigned to a 4 week YOCAS program and the assessments were made at baseline and post intervention. This study has some limitations they were taken from samples of only the white people, educated people and female breast cancer survivors. Future work is based on the different cancer types and comparing yoga with other treatments.

2.2 Yoga Classification and Prediction:

This study is all about presenting an hierarchical classification for yoga poses by implementing two important models such as DenseNet-201 and ResNet-50 and with the image augmentation techniques.Ghongane et al.(2022) takes the Yoga-82 dataset which has total of 82 different poses and was labeled into three levels 6, 20 and 82 classes all these images varied by the background color, indoor and outdoor images and more on. Data acquisition was performed in the yoga-82 dataset by addressing the drawbacks of the yoga

poses and then training, validating and testing the data files to increase the size of dataset size rotation has been done to maintain the balance. Finally the models were evaluated using Top1, 3 and 5 of the accuracy metrics where the DenseNet-201 accuracy was 95.45% and Resnet-50 accuracy was 94.67% and then the DenseNet-201 was the best performance model in this study. Due to the lack of training data and imbalance of the classes the model has been faced with overfitting. Also to improve the accuracy and balance the levels the dataset should be modified in future. This proposed study uses the Mediapipe library to classify the yoga poses with the given images and this study uses the models like VGG16 and VGG19 to improve the yoga posture classification. This study implemented the previous paper drawback of training data and classes. The VGG19 model consists of 19 deep layers based on the different indoor environments the dataset was created and to perform the correct postures the system identifies the relevant yoga training information. The VGG19 is the best fit model with an accuracy of 99.05%.. Anuradha et al.(2024) tells the limitations that even if the model has been performed well the quality and quantity of the dataset can still be improved. The future work can improve this technique with more deeper models such as ResNet50 and Inceptionv3. This study examines the Internet of things to monitor the yoga postures to get the yoga posture correctly by avoiding any incorrect postures. The novelty would be the quality and quantity of the dataset has been improved well. Pal et al.(2023) used datasets from different websites and then to enhance and detect the yoga postures through some machine learning approaches. By providing the automatic guidance, classifying, analyzing and monitoring yoga was the primary goal of this project. This study uses the data integration models which only focus on processing and analyzing the data from various sources which helps to preprocess and then mining the value of large and distributed datasets and then the accuracy scored around 92.34%. In this study the limitations that the researcher are telling is that the yoga is detected and monitored not in the effective way the future work could be implementing different kinds of models into the yoga pose detection. The proposed study examines the yoga pose identification provided by the voice feedback which helps to perform the posture correctly and reduce the risk of injury. The yoga 82 dataset used to find the posture correctly by using various images they wear scaled, normalized and converted to bitmap format. Kadam et al.(2024) to point the perfect yoga posture blazepose model have been used from machine learning which detected 33 keypoints and to check whether pose is correct or wrong K-Nearest Neighbors (KNN) have been used to classify them. In this study the metrics were specifically calculated for five different yoga poses they are tree, chair, warrior-1,2 and 3 which calculates the accuracy, precision, recall and F1-score. Author was faced with some challenge as there was some complexity in poses for the respective key points which has been making the classification tough. The future work that they mentioned is increasing the number of yoga poses and then refine the model. The proposed study explains that the system detects yoga posture and giving real time feedback based on the posture correction based on the key joints it will identify the key points which has been making the classification easy based on the predefined criteria almost 2500 images have been separated into seven different yoga poses Utkatasana, Bhujangasana, Adho mukha svanasana, salamba sarvangasana, Trikonasana, Vrikshasana and Virabhadrasana. Jadhav et al.(2023) have split the dataset into 70% and 30% for training and testing by resizing the images by 300x300 pixels. A deep learning based model, a modified version of Tensorflow called MoveNet model has been used in this project with 17 key body points. The Movement framework uses the mobile net V2 as the feature extraction also for performance they have used the MoveNet lightning and for accuracy they have used the MoveNet thunder with rate of 99.29% and 99.88% for training and testing respectively. The drawbacks that the author feels about this study is based on the clothing, lighting conditions and camera angles the systems performance have been affected and the future could be expanding the dataset which

includes more number of asanas also developing an mobile application. Mahindru et al.(2024) based on the people's mood, position detection, estimation and correction the systems implements and recommends the yoga estimation by using the advanced machine learning algorithms and computer vision methods the images consists of yoga poses which was associated with the mood categories and the images were resized into 224x224 pixels to match the predefined image size. Techniques such as data augmentation were used to increase the dataset size. The CNN pretrained models were used such as ResNet, VGG, MobileNet and efficientNet where the efficientNet is performed best among all the CNN models with an accuracy of 77.94%. To achieve the alignment and pose correction the researchers have faced some challenges. Moreover, they have been mentioned to expand to the maximum potential for the yoga pose correction module.

2.3 Image Prediction Using Inception-V3 and Transfer Learning:

This study explains about the classification of beef images which has been categorized into three parts mainly fresh, half fresh and spoiled it uses the InceptionV3 which is a pre trained model also uses the meat freshness image dataset which was sourced from kaggle and some where scraped from the web images consists of 1815 images for training and testing the data was preprocessed by augmenting the images and in the InceptionV3 model they were adding the layers using ReLU and Softmax functions also the adam optimizer was used and the model is trained with 50 epochs to reduce the covariance shifts several conv2d layers have been used and this model have reached an accuracy, precision, recall of 87.14% and F1-score of 88.28% respectively. The author falls back in identifying some limitations mainly to get more images in the dataset and in future the author suggests implementing different architectures to compare the performance and also additional features such as text and color for better accuracy. Rukhsar et al.(2022) focuses on developing the InceptionV3 model to classify the rice leaf diseases the dataset was sourced from kaggle and totally there were 2684 images for training and 671 for testing. Compared to the previous paper the challenges faced by the author were solved in this paper by improving the number of images. This study explains about the diseases which were classified into four types namely brown spot, Hispa, Leaf Blast and healthy leaves. To classify these types the author used pre pre-trained InceptionV3 model and transfer learning for feature extraction totally the model was trained for 40 epochs. Totally the accuracy lies above 90% following the precision, recall, F1-score for each of the types includes 76%, 85% and 92% respectively. The author faces some challenges to get a larger dataset and many types of rice diseases. The future work could be extended to other CNN models and compare them to improve the classification. Traditional methods like HOG and SIFT prove to be generalized to less accuracy. Sunday et al.(2023) examined the inception-v3 model to classify the fruit image for smart agriculture and harvesting robots using a modified algorithm for feature extraction to get the high accuracy compared to traditional methods. Compared to the previous paper this study has six types of fruits such as apples, bananas, kiwifruit, mangoes, citrus and pears the image pixels are in 299x299. Also by freezing certain layers and retraining the classifier the InceptionV3 model has been applied and implemented a softmax to classify the images into one of the 6 fruit images using 96 layers. The accuracy and loss values are 98.5% and 0.11 beyond 3000 iterations. The drawback is this study classifies only six types of fruit images which automatically reduces the effectiveness of fruit categories and it can expand the design systems to learn new fruit categories. This study examines the sensor set which can lower the injury while doing yoga, gym, swimming, golf by estimating the lower-limb joint angle with which the IMU set can be attached in every body. To address this transfer learning approach have been implemented in LSTM neural network to improve the model for new activities and

to minimize the new activity data. The dataset has five activities yoga, golf, swimming, badminton and dance for training and validation LOSO cross validation have been used. Methodology involved has two models LSTM and SVM with high accuracy of 92% and LSTM was the best fit model which incorporates the pre-trained source and transfer learning process. The author falls back in lower limb joint angle estimation which needs further testing across more various activities. In future it can extend new activities and can reduce the number of sensors without compromising the accuracy. In India the rice cultivation is more because of the significant reliance on rice. Half of the populations in India will have rice on a daily basis which then harms the crop with biotic and abiotic stress. The author Baskar et al.(2024) used deep learning techniques such as InceptionV3 and VGG16 to diagnose the paddy leaf disease to classify and identify what kind of disease does it occur. The dataset has a total of 800 images with 160 image perclass there are totally 5 classes namely bacterial blight, blast, tungro, brown spot, healthy leaf. The model was trained to InceptionV3 and VGG16 to handle the complex image classification tasks and the strong and depth baseline for the classification tasks. The best models are InceptionV3 and VGG16 with their accuracy of 97% when compared to the MobileNet and ResNet models. The author faces challenges in approaching the complexity and training time and it has a really low number of images to classify. The future work could be expanded to explore more techniques to reduce the computational overhead.

3.Research Methodology

This machine learning research will use the transfer learning architectural model from InceptionV3. The fig (1) shows the process of this study based on the yoga images. CROSS-INDUSTRY STANDARD PROCESS FOR DATA MINING methodology have been used in this study starting from understanding the data by collecting the data into 38 classes and pre-process then the data preparation by transforming the data into 80,20 split and in the modelling phase the mediapipe is used for pose detection, angle calculation and training the InceptionV3 model. Furthermore, evaluating the predicted yoga poses from skeleton images by calculating the performance metrics with some visualizations and the final stage is the deployment phase where the trained model used to predict the yoga poses. The first stage is data is pre-processed by augmenting the image and then converting the real time image to skeleton images to predict and get the metrics accurately. Followed by, the data is split into training and validation the data is trained to recognize the images. Then the training results are added to the InceptionV3 algorithm which is a transfer learning approach to improve the model performance. Then the data is validated and evaluated in the pre-trained new model. The new model is used to test whether the performance and accuracy accurately results from the model or not.

3.1 Data Collection:

Data collection refers to the process of obtaining, measuring, and analyzing data from various sources in order to form conclusions. Yoga positions were collected from numerous websites and organized into classes. To improve the model's accuracy, the collected photos were turned into a skeleton dataset using the OpenCV Python package.The data has been divided into 38 classes, with each class containing more than 150 images. The classes were, Adho mukha svanasana, Adho mukha vrksasana, Ananda balasana, Anantasana, Anjaneyasana, Ardha chandrasana, Ardha matsyendrasana, Bakasana, Bhekasana, Bhujangasana, Chaturanga dandasana, Dandasana, Dhanurasana, Halasana, Hanumanasana, Krounchasana, Natarajasana, Padmasana, Paripurna navasana, Parivrtta trikonasana, Paschimottanasana,

Prasarita padottanasana, Purvottanasana, Rajakapotasana, Savasana, Setu bandha sarvangasana, Tulasana, Urdhva dhanurasana, Urdhva prasarita eka padasana, Ustrasana, Utkatasana, Uttana shishosana, Uttanasana, Utthita hasta padangustasana, Vajrasana, Virabhadrasana ii, Vriksasana and Vrischikasana. The skeleton dataset was used by the Pre-Trained InceptionV3 model to classify yoga positions using a transfer learning strategy.

3.2 Data Preprocessing:

Once the data is collected the data is pre-processed and the existing asanas are augmented as right, left, rotate, up and down which are then used to expand the size and quality of the training set. Then the augmented real time images are converted as the skeleton image to find the accurate metrics by adding the data points using the deep learning approach which helps to classify what is that asana with the respective names.

3.3 Generating Skeleton Key Frames:

To transform the input image as the skeleton key frames, pose estimation the OpenCV python tool was used to predict and classify what type of asana it is. Media pipe are also used to estimate the skeleton images where the skeleton joints are connected using Media pipe it converts the normal coordinates to the pixel coordinates to draw lines and circles at the specified landmarks.

3.4 Model Training and Optimization:

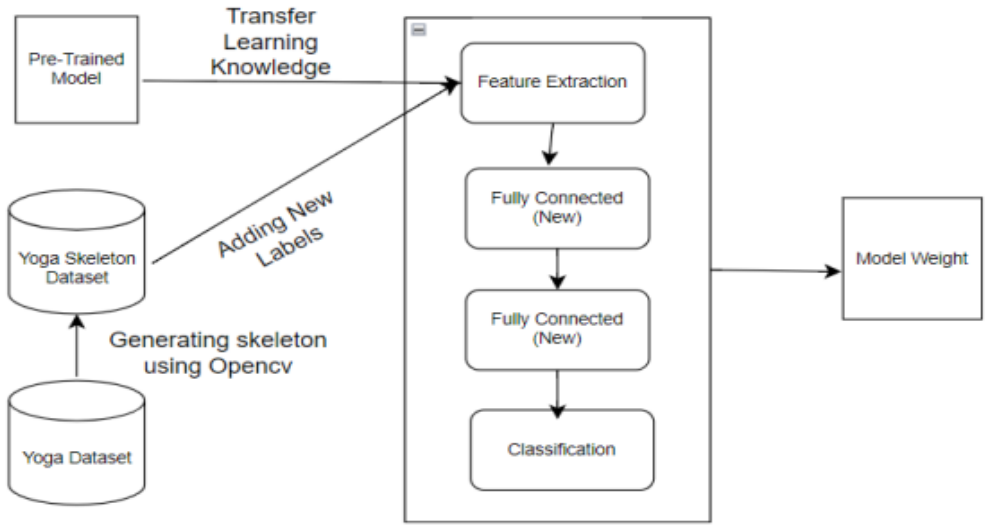

Figure 1 Transfer Learning architecture

Fig (1) shows how the transfer learning have been used to help in this study after loading the data and pre-processing it the features are extracted from pre-trained InceptionV3 model and it is connected to the new customized layers based on the classification task with each of the 38 classes and then checks the model weight. The dataset was divided into two parts training and testing, 80% of the data was utilized for training and 20% was used for testing. The model used in this study is InceptionV3 were trained on the processed dataset which helps to classify for the new tasks. Initially, 38 classes have been used to load the InceptionV3 model without the fully connected layers to allow the flexibility to allow new dataset by defining the model input shape as (224x224x3) to match the networks. For fine tuning the model the last

15 layers are pre-trained for the final layers where this is the common technique to adapt the new data. For feature extraction custom layers are added on top including normalization and stabilization with ReLU (Rectified Linear Unit) activation. Finally, 38 classes have been used to activate the softmax to classify the probabilities for each and every class and the model has been trained using the Adam Optimizer with loss function and accuracy as a metric which is suitable for multiple class classification tasks.

3.5 Pose Detection:

For detecting the pose an open source platform Mediapipe framework has been used for postures evaluation**.** Pose detection has been used to recognise the human body joints and skeleton in a photo. By using the Mediapipe pre-trained posture detection technique the 33 human body landmarks fig (2) from nose, shoulders, elbows to wrists were detected in real time. The landmarks have been employed in a variety of applications which includes pose detection, gesture detection and motion tracking. The mediapipe framework first extracts the joints from the image and then connects adjacent joints which are known as the edges using the draw function. The draw function is used to draw the lines connecting three specific joints to visualize the angles between them and understand the posture better.

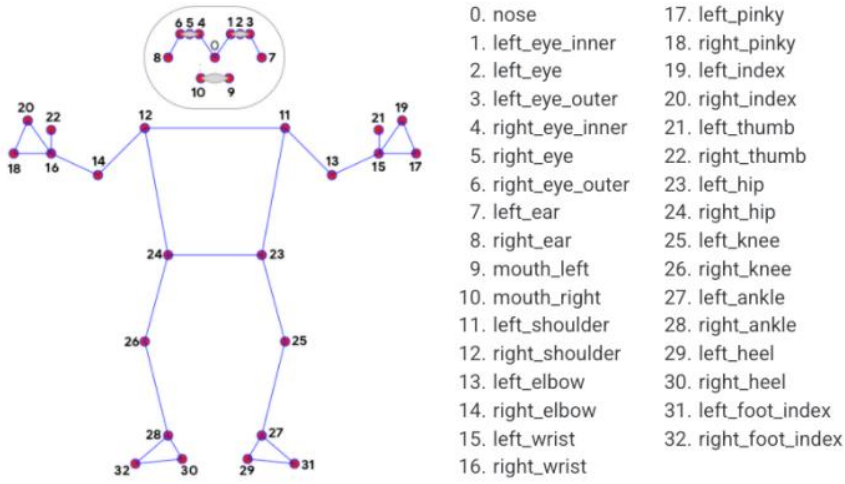

Figure 2 Mediapipe Model (Sunney et al.(2022))

3.6 Yoga Pose Prediction:

After the model is trained the real input image is converted as the skeleton image and then the model predicts the class of the given yoga asana and visualizes the result. For each of the images retrieved the model predicts the class label and it is determined.

3.7 Model Evaluation:

The model's performance has been evaluated by using the test data and finding out the accuracy, precision, recall, f1 score and also a graph for all these metrics. Correlation heatmap is also used to represent the data how well the machine learning model is set to the test data. Finally, a classification report shows how much the metrics have been calculated for

each of the asanas. These metrics provide a holistic view of the model's performance which considers the positive and negative classes.

4.Design Specification

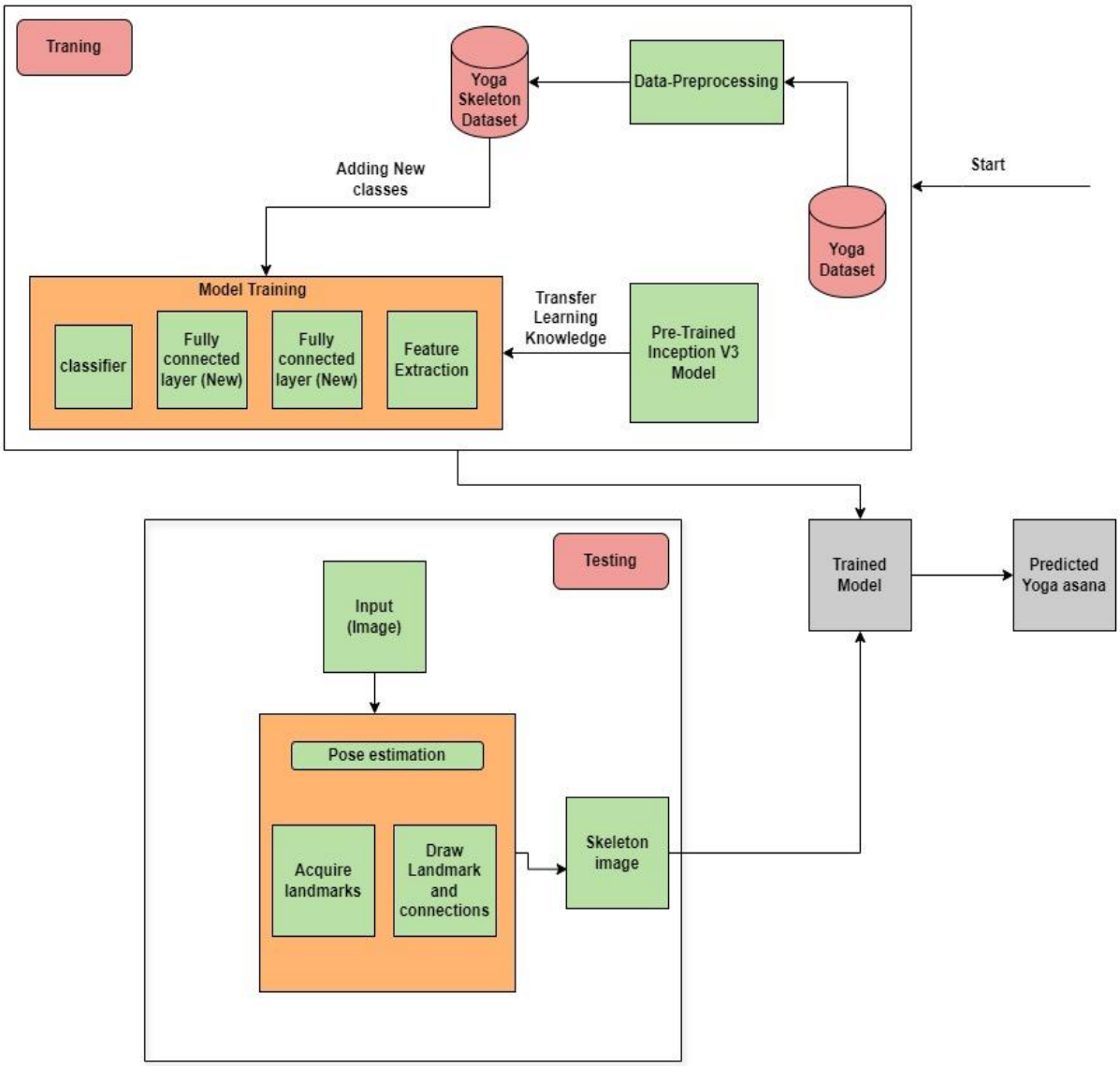

Figure 3 Architectural diagram of Yoga pose prediction

The purpose of this study is to improve the yoga pose prediction using the InceptionV3 algorithm and transfer learning model. The yoga pose prediction architectural framework in Fig 3 includes the data pre-processing, data augmenting, training the model using the pre-trained inception-v3 model then predicting the yoga asana with their respective names as a skeleton format. Firstly, taking the image of the person who is doing the yoga asana then the image is passes through initial convolutions and pooling layers to detect the basic features of them, then each of the inception modules captured through different features and scales while training the model the network help to learn better about the joint positions and then the extracted features are combined to make final prediction. Finally, the output represents the raw image yoga as the skeleton format to provide the accurate metrics while predicting the asanas.

4.1 Transfer Learning Architecture with InceptionV3:

Transfer learning is a technique which is used in machine learning where in one model there will be a particular task which has been used, the used model is again reused for the second model as the starting point. In simple terms the knowledge which is used in one task helps to solve the other related task. The InceptionV3 model is used to build several convolutions with different filter sizes for eg. 1x1 Convolutions: which helps to reduce the parameters and introduces the non-linearity. The 3x3 and 5x5 convolution helps to capture median and large features of the input whereas the pooling layers reduces the dimensions and helps with small translations. The layers which are explained in Fig 4 which are used to prevent the gradient flow are known as the auxiliary classifiers. Some benefits are there while using InceptionV3 are the number of parameters and computational cost which helps to reduce them and it is so flexible to fine tune on the different datasets.

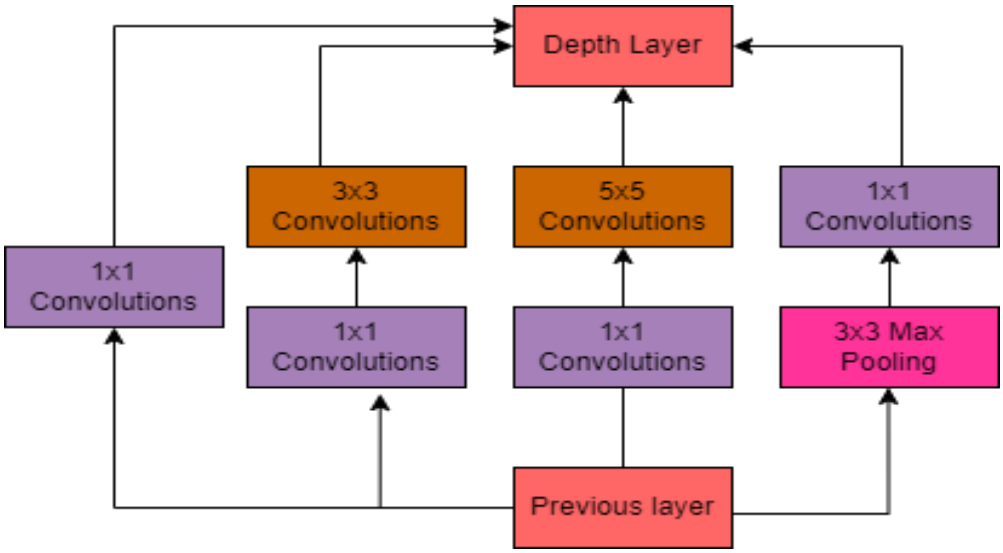

Figure 4 InceptionV3 architecture

5.Implementation

5.1 Tools & Libraries Used:

A personal fitness for yoga using InceptionV3 and transfer learning helps to predict the yoga asana accurately with the respective names by giving the raw image and predicts the asana through skeleton format. This section explains about the tools and libraries that were employed in this project. Jupyter notebook have been used to implement the code which helps to provide a wide range of vital resources for python developers. Jupyter notebook is a compiler where it helps to run the python code numpy, pandas, OpenCV, matplotlib and other additional libraries used to run the code. Altogether work to predict the yoga asana.

5.2 Dataset Used:

For prediction, the model used the yoga posture dataset. The Yoga Posture Dataset was used to implement the algorithms. That data set contains 38 unique yoga positions. The collection

initially contains human pictures, which were then converted to skeleton format for optimal accuracy. Each class consists of 108-273 pictures, every picture gathered from a different resource and stored as a single dataset in Kaggle. The dataset contains a total of 6702 images.

5.3 Data Pre-Processing:

Once the data is loaded some pre-processing steps have been implemented. Firstly, augmenting the 38 asanas to get more number of images then to detect the various landmarks and identify the position in image "findposition" functions are used and the "findjoints" function helps to draw lines between the specified joints to visualize the connections. Using the Mediapipe library it checks all the images one by one to detect the pose landmarks with a certain confidence threshold. If the landmarks are detected then it creates a white blank image where it draws the pose skeleton. This is used to visualize pose estimation on a clear background for future analysis. After loading the images each image is resized into 224x224 pixels which then helps to predict the image with their respective names.

5.4 Feature Extraction:

The dense layers are added with the ReLU activation which helps to extract the high level features specific to the prediction task and then the features are extracted with the pre-trained InceptionV3 model. Batch normalization helps to accelerate the training process to ensure that the training features are fed into the dense layers which are normalized. Starting with the basic edges in the early layers and then progressing to more complex patterns which used to object the parts in deeper layers for the image prediction.

$$F\hat{} = \text{BatchNorm}(\text{ReLU}(\text{InceptionV3}(X)))$$

From the above formula it defines that the X is denoted as the input image, followed by the InceptionV3 (X) is denoted as the feature maps, ReLU applies the activation function and BatchNorm normalizes the resulting feature maps. This formula captures the feature extraction from the input image and the activating those features and then normalizing them.

5.5 Feature Selection:

This code freezes the last 15 layers of the model by making these last layers to be trainable. This is because there are large datasets to fine tune the last few layers in order to adapt the yoga pose prediction. To reduce the dimensionality of the features output by adding 128 units as the dense layers for the feature selection which focuses on the most important features needed for the prediction. Softmax activation selects in the final dense layer by classifying the input into specified 38 classes which then extract the features into 38 classes.

$$y\hat{} = \text{Softmax}(\text{ReLU}(Wd \cdot F\hat{} + bd))$$

The formula explains the feature selection process by capturing the reduction in dimensionality which was classified from the feature extraction process. Firstly, F^ know as the normalized feature vector from the feature extraction process, followed by Wd and bd which are the weights and bias of the dense layers, ReLU is same which is the activation function. Finally, the Softmax is used for the 38 classes by providing the final prediction probabilities

5.6 Model Training:

Yoga poses have been collected from multiple websites and grouped into classes. The OpenCV Python package was used to transform the collected pictures into a skeleton dataset. The collection is divided into 38 classes, each with more than 150 images. Understanding of feature extraction knowledge obtained by using the pre-trained Inception-V3 model. Retraining the entire model's layers with new data will be avoided. Only remove the last layer and replace it with an output layer with the number of classes in the dataset. Model training accuracy is 96.03% using the skeleton dataset. By omitting the top layers and retaining just the last 15 layers to learn the features which allows the fine tuning some custom layers are then added on top which includes the batch normalization the dense layer with ReLU activation and the final layers of softmax activation to enhance the probability of classes in the dataset. This model is compiled using the Adam optimizer and the loss which is helpful for classification tasks also this model helps to specify the task with less training data and computational resources.

The final model training step used to specify the prediction task by loading it without the top layers and only used the last 15 layers with some required customization. Same as the last step, this final model training step also includes the ReLU activation, softmax for dense layers to predict the class probabilities based on the number of classes. The model is trained on the number of training and validation steps to calculate the dataset size and batch size the total layers used while training the model is 312 layers and during training, monitoring the validation accuracy and loss the epoch is 200 in order to learn the training data too well because of the large dataset also predict more yoga poses the model needs more epoch.

6.Evaluation & Results

The model is just an image processing model. First checking whether the images from the dataset are first predicting perfectly or not from fig (5) this is the raw image. The asana which they are predicting known as the Vrischikasana.

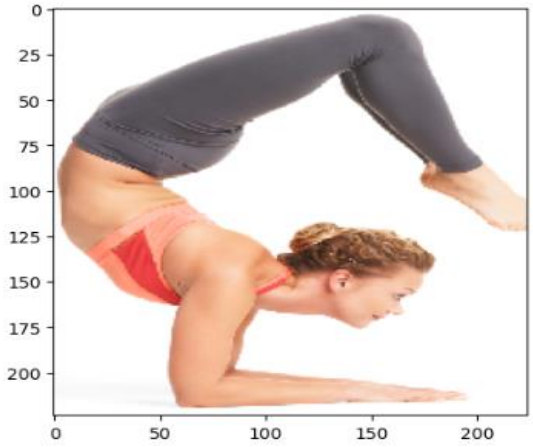

Figure 5 Vrischikasana

Then checking whether the skeleton image is predicted properly or not after training the data from fig (6) is the predicted image model with their respective names. This data was taken

from the skeleton dataset which changed the normal image to skeleton image while doing the pre-processing step.

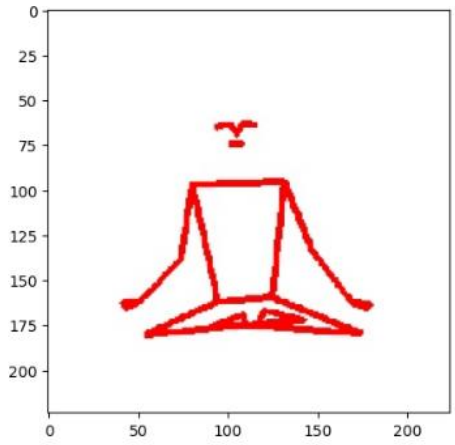

Figure 6 Predicted skeleton model

6.1 Model Prediction:

Using Mediapipe the human pose image helps to detect the InceptionV3 model by extracting the coordinates of the key landmarks by annoting these landmarks and then draw the connections between them to visualize the human pose. In fig (7) represents the path where all the raw yoga pose images which have been performed by the practitioner are stored and it helps to give an output of the skeleton image with the respective names.

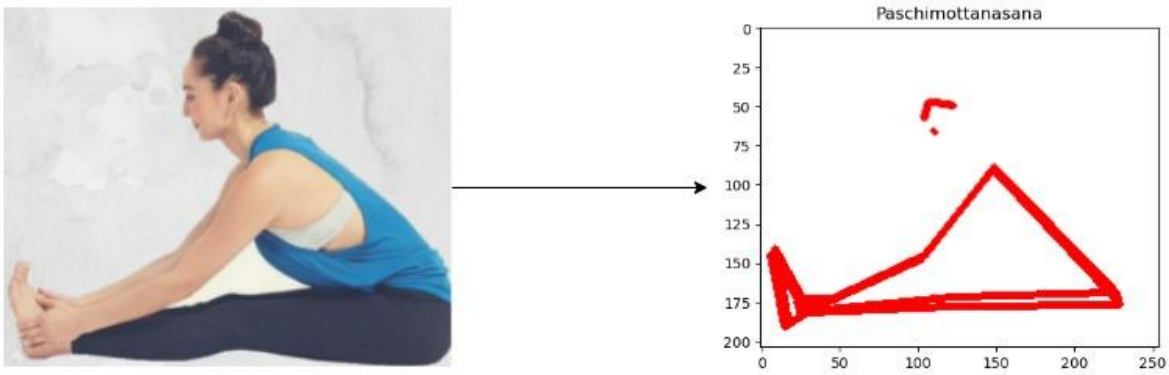

Figure 7 Final Output with the respective name

6.2 Performance Analysis:

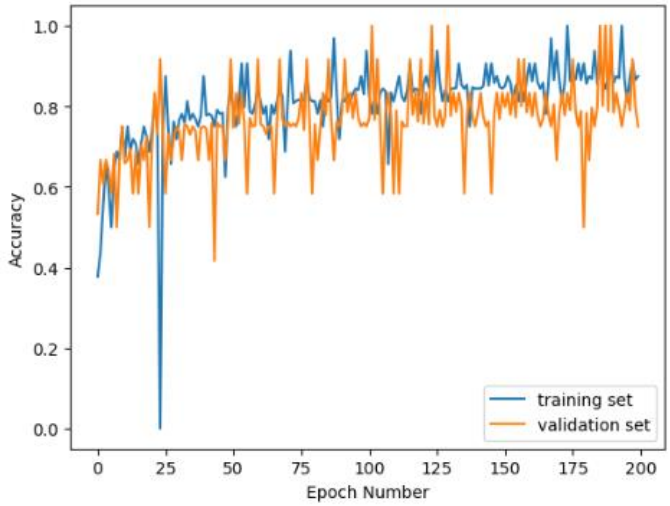

Figure 8 Model Accuracy

Evaluation Metrics used to evaluate the model's performance certain evaluation metrics were discussed below. The model used in this work is InceptionV3. This model's accuracy fig (8) was higher than that of other models such as Spatial-Temporal Fusion Residual Network (STF-ResNet) Wanjun et al.(2023) and AdaBoost Kundu et al.(2023) by Rajendran et al.(2024). The training accuracy of the model is 96.03%. And model validation accuracy of the model is 90.95%. The loss fig (9) of 10% accuracy generated by Ardha Matsyendrasana, Bhekasana, Kronchasana, and Vajrasana in the Model Training. The training loss of the model was 0.3206, and the testing loss was 0.7347. As shown in Fig. (10) Model's Performance based on Loss.

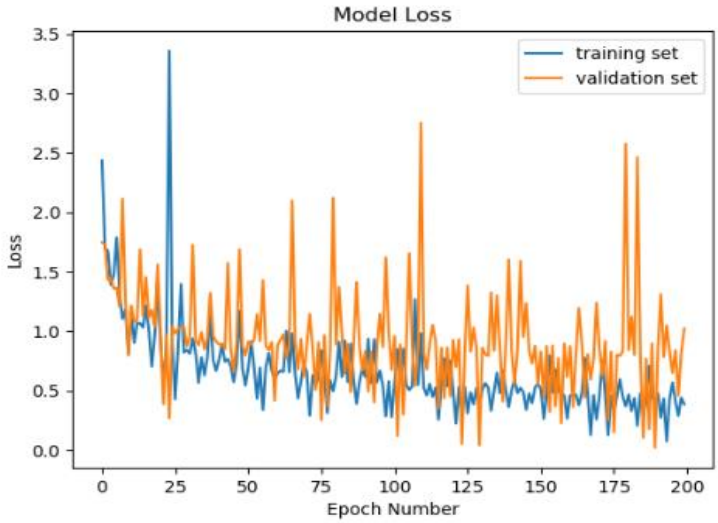

Figure 9 Model Loss

6.3 Comparative Analysis:

Paper	Model Used	Dataset Used	Accuracy
Classification of yoga pose using integration of deep learning and machine learning techniques	AdaBoost	ImageNet dataset taken from many resources.	93%
Yoga action recognition based on STF-ResNet.	STF-ResNet	Custom Yoga dataset	95%

Figure 10 Comparative Analysis Table

This study was compared with various existing works and the proposed work as well. A quick overview of this referred paper shows in fig(10) which demonstrates how well the accuracy score is high compared to other different machine learning and deep learning models such as AdaBoost, STF-ResNet and VGG model.

6.4 Confusion Matrix:

A table known as a confusion matrix is frequently used to represent how well a machine learning model performs on a set of test data for which the true values are known. A number of evaluation measures for the machine learning model, including accuracy, precision, recall, F1 score, and others, are computed using the confusion matrix. These indicators help in understanding the model's performance and identifying potential areas for development. By computing model accuracy for each class using a confusion matrix with the entire dataset as input

6.5 Classification Report:

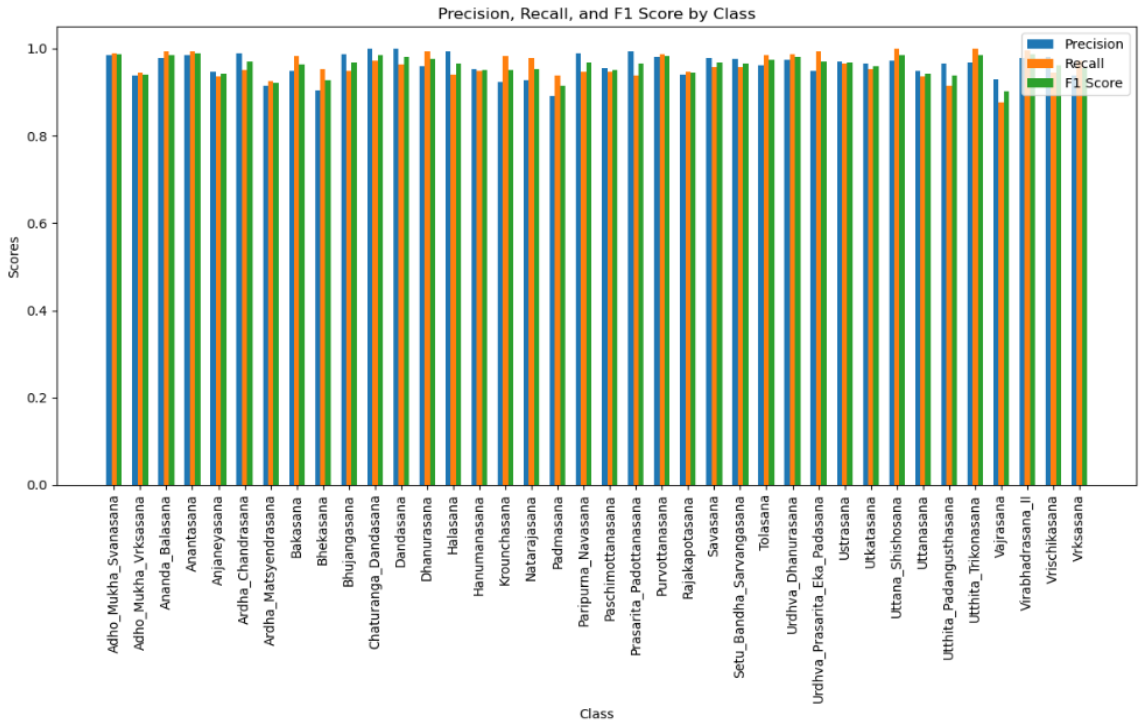

Figure 12 Bar graph of Metrics

Classification report shown in fig (13) has the Precision, Recall, F1-score, support score of the model. To see the precision, recall and F1-score based on the 38 classes a bar graph is shown in fig (12).

- **F1 score:** It is an evaluation metric for analysing the model's performance. It used precision and recall. The maximum f1 score is 1 and the minimum f1 score is 0.

- **Precision:** It checks how well the model can predict the class. It is defined as the number of true positives divided by the total number of positive predictions.

- **Recall:** Recall measures the proportion of actual positive instances (true positives) that were correctly identified by the model as positive, out of all the positive instances that exist in the dataset.

Class	Precision	Recall	F1-Score	Support
Adho_Mukha_Svanasana	0.98	0.99	0.99	187
Adho_Mukha_Vrksasana	0.94	0.94	0.94	127
Ananda_Balasana	0.98	0.99	0.99	133
Anantasana	0.99	0.99	0.99	134
Anjaneyasana	0.95	0.94	0.94	174
Ardha Chandrasana	0.99	0.95	0.97	185
Ardha Matsyendrasana	0.92	0.93	0.92	188
Bakasana	0.95	0.98	0.96	167
Bhekasana	0.90	0.95	0.93	128
Bhujangasana	0.99	0.95	0.97	259
Chaturanga Dandasana	1.00	0.97	0.99	179
Dandasana	1.00	0.96	0.98	108
Dhanurasana	0.96	0.99	0.98	143
Halasana	0.99	0.94	0.97	166
Hanumanasana	0.95	0.95	0.95	173
Krounchasana	0.92	0.98	0.95	110
Natarajasana	0.93	0.98	0.95	236
Padmasana	0.89	0.94	0.91	273
Paripurna Navasana	0.99	0.95	0.97	189
Paschimottanasana	0.95	0.95	0.95	133
Prasarita Padottanasana	0.99	0.94	0.97	180
Purvottanasana	0.98	0.99	0.98	154
Rajakapotasana	0.94	0.95	0.99	151
Savasana	0.98	0.96	0.97	185
SetuBandhaSarvangasana	0.98	0.96	0.97	166
Tolasana	0.96	0.99	0.97	203
Urdhva Dhanurasana	0.97	0.99	0.98	148
Urdhva Prasarita Eka Padasana	0.95	0.99	0.97	133
Ustrasana	0.97	0.97	0.97	230
Utkatasana	0.97	0.95	0.96	213
Uttana Shishosana	0.97	1.00	0.99	136
Uttanasana	0.95	0.94	0.94	222
Utthita Padangusthasana	0.97	0.91	0.94	151
Utthita Trikonasana	0.97	1.00	0.98	277
Vajrasana	0.93	0.88	0.90	211
Virabhadrasana II	098	0.99	0.99	192
Vrischikasana	0.98	0.94	0.96	162
Vrksasana	0.94	0.97	0.95	185
Accuracy			0.96	6691
Macro Avg	0.96	0.96	0.96	6691
Weighted Avg	0.96	0.96	0.96	6691

Figure 13 Classification Report of 38 asanas

7.Conclusion and Future Work

The objective of the proposed system is to predict the correct yoga posture with their respective names. The system helps to identify and correct 38 different yoga postures with an accuracy of 96.03%, this level of accuracy is critical as it performs the yoga posture correctly. Using a transfer learning approach, the InceptionV3 model is used to train and test the dataset. One of the main aspects while predicting the yoga pose by using Mediapipe which helps to detect the human the key points from the raw images and to enhance the pose prediction the raw images have been changed to skeleton format to improve the model as well. The features from the image have been extracted, and angles have been calculated in order to adjust the posture. The study assists in predicting human posture and provides tips to improve posture. To improve the model performance as well this research highlights the importance of pre-processing and data augmentation as well which also helps to set a benchmark for future research. Finally, evaluation metrics have been included such as accuracy, precision, recall, F1-score which helps to assist the model's performance. As a result, the proposed system helps the user acquire all of yoga's benefits.

There are more than 80 different yoga postures; however, the proposed system is limited to just 38, future research can expand the dataset for in a broader range which then helps to annotate more asanas and give good accuracy. The proposed dataset can be increased by including the key points for the necessary yoga poses. The system only recognises a single person; it can be extended to recognise multiple people in the image which can help for predicting the group yoga poses. Can develop a user system where they can monitor and give feedback in real time with an audio-based feedback by setting the range as if the user is not posing properly then the user needs to perform the pose again. Also, the future work can include by using the wearable device which can be used with the help of smartphones or motion sensors which then helps to lead more accurate posture correction.

References

Akdeniz, Ş. and Kaştan, Ö., 2023. Perceived benefit of yoga among adults who have practiced yoga for a long time: a qualitative study. *BioPsychoSocial Medicine*, *17*(1), p.19.

Mahindru, A., Patil, P. and Agrawal, V., 2023. Role of physical activity on mental health and well-being: A review. *Cureus*, *15*(1).

Patel, R. and Veidlinger, D., 2023. Exploring the benefits of yoga for mental and physical health during the COVID-19 pandemic. *Religions*, *14*(4), p.538.

Telles, S., Sharma, S.K., Gandharva, K., Prasoon, K. and Balkrishna, A., 2023. Yoga Practice and Choices of Foods, Physical Activity, and Leisure: A Convenience Sampling Survey from India. *International Journal of Yoga*, *16*(1), pp.20-26.

Lin, P.J., Altman, B.J., Gilmore, N.J., Loh, K.P., Dunne, R.F., Bautista, J., Fung, C., Janelsins, M.C., Peppone, L.J., Melnik, M.K. and Gococo, K.O., 2023. Effect of yoga and mediational influence of fatigue on walking, physical activity, and quality of life among cancer survivors. *Journal of the National Comprehensive Cancer Network*, *21*(2), pp.153-162.

Ghongane, Aishwarya (2022) Hierarchical Classification of Yoga Poses using Deep Learning Techniques. Masters thesis, Dublin, National College of Ireland.

Anuradha, T., Krishnamoorthy, N., Kumar, C.P., Prasad, L.N., Chunduru, A. and Moorthy, U., 2024. A Method for Specifying Yoga Poses Based on Deep Learning, Utilizing OpenCV and Media Pipe Technologies. *Scalable Computing: Practice and Experience*, *25*(2), pp.739-750.

Pal, R., Adhikari, D., Heyat, M.B.B., Ullah, I. and You, Z., 2023. Yoga meets intelligent internet of things: recent challenges and future directions. *Bioengineering*, *10*(4), p.459.

Kadam, P., Kadam, S., Bidwe, R., Shinde, N., Ginnare, N. and Kesari, N., 2024. Smart Yoga: Machine Learning Approaches for Real-Time Pose Recognition and Feedback. *International Journal of Computing and Digital Systems*, *16*(1), pp.1-10.

Jadhav, R., Ligde, V., Malpani, R., Mane, P. and Borkar, S., 2023. Aasna: Kinematic Yoga Posture Detection And Correction System Using CNN. In *ITM Web of Conferences* (Vol. 56, p. 05007). EDP Sciences.

Duppala, V.R., Marepalli, H.Y., Jain, K., Anusha, K., Thangavel, S.K., Kumar, B.S., Bindu, A., Satish, L. and Sekar, J., 2024, April. Aatma Yoga: Automation of Yoga Pose Recognition and Recommendation using Deep Learning. In *2024 International Conference on Inventive Computation Technologies (ICICT)* (pp. 1315-1322). IEEE.

Upadhyay, S.K., 2022, May. Deep transfer learning-based rice leaves disease diagnosis and classification model using inceptionv3. In *2022 international conference on computational intelligence and sustainable engineering solutions (CISES)* (pp. 493-499). IEEE.

Sunday, S.E., Ji, R., Abdalla, A.N. and Bian, H., 2023, October. Fruit Image Classification using the Inception-V3 Deep Learning Model. In *2023 International Conference on the Cognitive Computing and Complex Data (ICCD)* (pp. 227-230). IEEE.

Li, J., Zhu, K., Li, D., Kang, P. and Shull, P.B., 2024. 3d knee and hip angle estimation with reduced wearable imus via transfer learning during yoga, golf, swimming, badminton, and dance. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*.

Sowmiya, B., Saminathan, K. and Devi, M.C., An Ensemble of Transfer Learning based InceptionV3 and VGG16 Models for Paddy Leaf Disease Classification.

Sunney, Jothika (2022) Real-Time Yoga Pose Detection using Machine Learning Algorithm. Masters thesis, Dublin, National College of Ireland.

Kundu, K. and Goswami, A., 2023, March. Classification of Yoga Poses Using Integration of Deep Learning and Machine Learning Techniques. In *Proceedings of International Conference on Recent Trends in Computing: ICRTC 2022* (pp. 417-428). Singapore: Springer Nature Singapore.

Wanjun, Y., Chong, C. and Rui, C., 2023, January. Yoga action recognition based on STF-ResNet. In *2023 IEEE 3rd International Conference on Power, Electronics and Computer Applications (ICPECA)* (pp. 556-560). IEEE.