Physical Exercise Pose Detection using BlazePose and Machine Learning Framework

MSc Research Project
Data Analytics

Aarthi Rajendran
Student ID: 22224947

School of Computing
National College of Ireland

Supervisor: Ahmed Makki

National College of Ireland

MSc Project Submission Sheet

School of Computing

Student Name: ....Aarthi Rajendran...

Student ID: ....22224947...

Programme:.........Research Project...................................... **Year:** ......2024...................

Module: ......MSc Data Analytics...

Supervisor: ....Ahmed Makki...

Submission Due Date: ....12/08/2024..

Project Title: .... Physical Exercise Pose Detection using BlazePose and Machine Learning Framework ....

Word Count: ....5794........................... **Page Count**.................16..

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: ....Aarthi Rajendran..

Date: ....11th August 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):	

Physical Exercise Pose Detection using BlazePose and Machine Learning Framework

Aarthi Rajendran
22224947

Abstract

Pose detection involves identification of specific body postures and movements. This approach can be applied to various exercise activities, providing insights with widespread applications in fitness, sports, healthcare, and human-computer interaction. This research proposes automated pose detection and classification algorithms to automate the identification of exercise activities, providing a scalable and effective solution for real-time monitoring and analysis in various factors such as enhancing fitness tracking, improving sports performance engaging more people into the physical activity. This research presents a comprehensive approach to detecting human poses and identifying specific exercise activities using machine learning techniques. By utilizing MediaPipe for pose detection and the classification algorithm able to detect the key points accurately on the human body and classify them into predefined exercise categories. The proposed framework enhances real-time monitoring of human exercise poses by integrating pose detection models with classification algorithms. Using pose detection models like MediaPipe and BlazePose, which extract body positioning features in the form of coordinates (x, y, and z). These coordinates are subsequently used to generate custom geometric features for model training. A large-scale dataset from UCF 101 consists of realistic action videos having 101 different action categories. This data set is an extension of UCF50 data set which includes 50 action categories and features 13320 videos across the 101 action categories, where 7 groups are taken for our research such as pushups, pullups, Body weight squats, bench press, and jumping jacks. Using three algorithms to classify the exercise activities. Among these, XGBoost is identified as the best model, achieving an accuracy of 94%. followed by GBM with an accuracy of 93% and Distributed Random Forest (DRF) with an accuracy of 91%. This framework enhances real-time monitoring of human exercise poses, offering significant potential for applications in automated fitness tracking and sports analytics. This research aligns with the cutting edge in pose detection and classification, providing real time monitoring solution by using advanced algorithms and machine learning for high accuracy by making it suitable for various fitness applications. This research has potential for practical applications such as fitness tracking, and health monitoring. Future work could address the integration into the wearable devices to provide real time feedback and insights.

1 Introduction

Physical exercise pose detection has gained momentum post-pandemic as people increasingly prioritize their health and use the internet for seamless workouts. This rapidly evolving technology within computer vision has a wide range of applications across various domains. Recent advancements, such as the BlazePose model, have significantly improved the accuracy of analysing body movements. By leveraging video footage of exercise workouts and processing human skeleton data with various machine learning algorithms, pose detection can accurately predict physical activity poses. (Min, 2022a)Integrating pose detection with advanced machine learning techniques enhances the precision and scalability of fitness tracking and monitoring, offers more effective solutions for analyzing and improving exercise performance. Pose detection is a rapidly advancing area within computer vision, specializing in the identification and analysis of human body postures and movements. By enabling detailed monitoring and analysis of body movements, pose detection provides crucial insights into exercise performance, technique, and overall physical activity, thereby transforming our approach to fitness and health. In the realm of fitness, pose detection facilitates the automation of exercise identification and classification, significantly enhancing tracking capabilities and improving sports performance. Integrating pose detection with advanced machine learning algorithms further strengthens these benefits by offering more precise, efficient, and scalable solutions. The primary approach to enhance the pose detection accuracy and efficiency involves combining the MediaPipe model with BlazePose. MediaPipe's framework, known for its lightweight nature and fast processing capabilities, is particularly effective in real-time pose detection(Zhang et al., 2022). When integrated with BlazePose, which excels in detailed body landmark detection, this combination significantly improves the accuracy of identifying and classifying exercise poses. This aligns with the current research focus, which emphasizes leveraging advanced machine learning models to refine exercise pose detection. By utilizing these sophisticated models, the research aims to advance the precision and efficiency of pose detection systems, thereby offering more robust and scalable solutions for fitness tracking and analysis. In MediaPipe framework (Amrutha, Prabu and Paulose, 2021)the ability to integrate multiple tracking modalities includes human landmarks and hand movements into a single pipeline. This provides detailed body analysis which enhances accuracy and effectiveness of pose detection systems.
The research objective are derived as follows: -

1. Develop a pose detection solution which utilizes the 2D (Setiyadi *et al.*, 2022)human body pose estimation to improve the accuracy of detecting the exercise poses.
2. (Krishnan *et al.*, 2022)Implement a machine learning model to identify physical activity poses and predict the type of exercise a person is doing.
3. Widely analyse and implement by using keypoint detection and machine learning to improve both the speed and accuracy of pose detection.
4. Evaluate and analyze the keypoints for performance monitoring and enhance to improve the accuracy for best prediction.

The classification of exercise poses and pose estimation have profound implications in fitness tracking and healthcare fields, as it enables the monitoring of self-progression and enhance the mobility and improves the fitness outcomes. (Wang, 2024a)This research plays a vital role in developing a scalable, pose detection solution with 2D human body pose estimation where only the keypoint information in the 2D coordinate system is considered. By integrating machine learning techniques with pose detection models like MediaPipe and

BlazePose, this research aims to detect the human exercise pose and improve accuracy and efficiency. A training approach is implemented by involving keypoint detection and machine learning techniques, During the training phase model is trained and tested with the video dataset by converting into H2O frames and enhance the detection speed without compromising accuracy. The final outcome includes 33 keypoints representing the coordinates where human movements, especially all joint angles, are visible. These keypoints are effectively utilized in strength training exercises to monitor and improve performance. The coordinates of the detected keypoints are carefully analyzed, as the joint angles play a crucial role in evaluating movements and postures. This analysis is essential for ensuring the accuracy and effectiveness of the exercises, contributing to better strength training outcomes.
This research implements various classification models, including XgBoost, GBM, and DRF, and evaluates their performance on a video dataset to identify the most effective one. By analysing the evaluation results from all three models, the most frequent prediction is finalized as the best-fit model. This combined approach significantly enhances the ability to track and analyze exercise movements, providing valuable insights for fitness and healthcare applications.
This research is based on the UCF101 dataset.1. which contains specific types of exercises. Because the dataset only includes certain exercises, the model may struggle to accurately recognize other exercises that aren't part of this dataset. In addition to that model performance may varies due to the factors like different body types, lighting in the videos and the video quality. These variations can affect how well the model detects and classifies exercises, potentially impacts the overall accuracy and reliability of the system.
The research study is structured as follows. Section 2 consists of Related Work about human activity and exercise related pose detection methods. In Section 3 the methodology about the research is described. Design specification is followed in Section 4 whereas in Section 5 implementation steps are summarized. The Evaluation is in Section 6 described in detail with all the experiments conducted based on the results and followed with that Discussion section. Finally, section 7 concludes the research and discusses the potential future work.

2 Related Work

In this current technology, Human pose detection is the efficient way to track and understand the human movements and body positions. It focuses on monitoring and analysing how people position their bodies and moves during the workouts. This helps ensure exercise are done correctly, improves performance, and enhances overall workout effectiveness.

(Sajan and Anilkumar, no date)has proposed a yoga monitoring pose estimation technique that focuses on improving efficiency and real-time performance. The system is designed to identify any errors or deviations from the correct posture. Google's MediaPipe framework is highlighted as a state-of-the-art solution, combining deep learning models with optimized processing pipelines to achieve real-time pose estimation on both mobile and desktop platforms. However, MediaPipe operates in a 2D space, which can lead to inaccuracies in pose estimation due to the lack of depth perception.

(Angelini and Naqvi, 2019) has used the UCF101[1] dataset in their research to evaluate and compare the performance of their proposed Joint RGB-Pose based multimodal networks. By

[1]Human action recognition Dataset :https://www.crcv.ucf.edu/data/UCF101.php

using UCF101, the author aimed to demonstrate how their multimodal approach—combining RGB and pose-based features—improves recognition accuracy and processing time compared to existing methods that rely solely on pose data.

(Irfan and Abdul Muthalib, 2022)The author introduces a solution involving "Human Pose Estimation," a technology that uses webcams to track and analyze physical movements. This allows individuals to exercise at home effectively. In detail the author has explained how human pose estimation works by tracking the movements, optimal distance for webcam detection and also clearly noted that that the use of MediaPipe results in accurate and effective detection of fitness reps.

(Bazarevsky et al., 2020) has introduced BlazePose, a lightweight neural network designed for real-time human pose estimation on mobile devices. BlazePose efficiently identifies 33 body keypoints from a single person and operates at over 30 frames per second on a Pixel 2 phone. This technology is ideal for applications such as fitness tracking and sign language recognition. The main contributions are the development of a novel body pose tracking method and a streamlined neural network that combines heatmaps and regression to determine keypoint coordinates.

(Singh, Panthri and Venkateshwari, 2022)In this paper, the author describes a method for measuring human body parts using pose estimation techniques. They utilize contour detection and the MediaPipe model to identify and track semantic key points on the body. By applying edge detection methods with OpenCV, the system determines body boundaries for accurate measurement. The goal is to develop a smartphone application that captures a 2D image from a front view to estimate body measurements such as waist and chest size. This approach aims to enhance online shopping by allowing users to accurately measure their body dimensions using a standard smartphone camera, ultimately improving their shopping experience.

(Zhang *et al.*, 2022)researcher has introduced a method for detecting deep squat movements by combining YOLOv5, a modified target detection algorithm, with MediaPipe, a human pose estimation framework. This approach enables accurate tracking of squat exercises, measuring angles in the trunk, hips, and knees with over 96% accuracy while reducing false detections. And this helps for our research in order to demonstrate precise movement analysis, which could enhance the accuracy and effectiveness of tracking various fitness activities.

(Mroz *et al.*, 2021) has compared BlazePose and OpenPose for virtual motion assessments using ten clinically relevant movement videos. While OpenPose provides more accurate keypoint locations and was preferred for clinical assessments, BlazePose offers faster runtime and could be used for tasks like movement pre-screening or activity classification.

(Nunes, Faria and Peixoto, 2017)The author introduces a novel framework for real-time human activity recognition using minimal training data. It features activity-based key pose extraction and trains a Random Forest (RF) classifier, showing strong performance on the Cornell Activity Dataset (CAD-60). The standard RF achieved 81.73% precision and 79.01% recall, while the Differential Evolution-enhanced RF (DERF) slightly improved these metrics to 81.83% precision and 80.02% recall. The RF method excels in avoiding overfitting, offers fast training, is robust to noise, and ease of parallelization.

(Bukht and Jalal, 2024)The author has developed a human activity recognition system using XGBoost, demonstrating a 91% recognition rate which shows that achieved the highest performance, primarily due to its effective integration of preprocessing, feature extraction, and classification techniques. So, it is a good decision to integrate the XGBoost-based human activity recognition approach into the "physical activity Pose Estimation" research.

(Wang, 2024b)The researcher develops an intelligent long jump evaluation system that improves teaching and assessment by addressing issues such as reliance on teacher experience and lack of quantitative feedback. By utilizing this insight for future work, we can enhance this approach by providing actionable feedback and deriving data-driven insights into exercise techniques.

(Song and Chen, 2024)The author develops a 3D bone keypoint detection algorithm and a pose estimation-based counting algorithm that achieves high accuracy, with a 96% detection rate in various environments and notable improvements in pose estimation. This counting algorithm is highly required for the evaluation stage of the research paper on human exercise pose detection as it provides precise, real-time feedback on exercise movements, enabling accurate counting and analysis of exercise poses to improve training effectiveness and technique correction.

3 Methodology

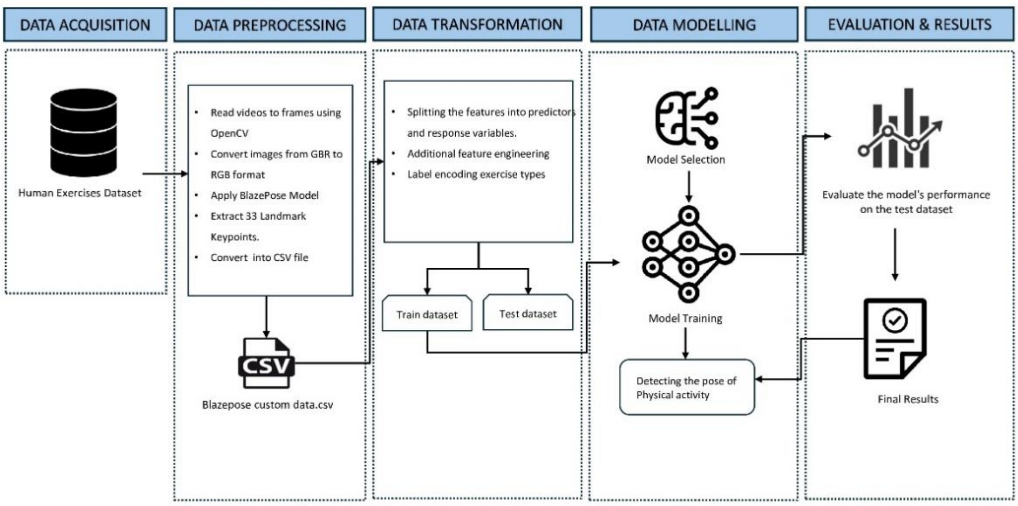

Figure.1 Methodology - Physical Exercise Pose Detection

The research methodology, depicted in Figure 1, provides a structured approach to this study, which consists of five steps such as Data Acquisition, Data Preprocessing, Data Transformation, Data Modelling and Evaluation. In the first step Data gathering process is performed where the dataset contains different human activity poses collected for this research. The dataset was created for human pose estimation using human action recognition from UCF101, since this dataset gives the high accuracy on human action recognition where this video data.(Angelini and Naqvi, 2019) In this paper, the author utilized the UCF101 dataset, which contains 13,320 multi-target video samples spread across 101 different actions in diverse contextual environments.

The experimental results demonstrate good accuracy, speed, and efficiency in the proposed method for joint RGB-pose-based human action recognition, aimed at anomaly detection applications. For the purpose of this research, the dataset includes a focused subset within the Exercise category, comprising videos of various exercises, including Push Ups, Jumping Jacks, Pull Ups, Body Weight Squats, and Bench Press.
During the preprocessing stage, the videos are converted into H2O frames, and the dataset is cleaned to retain only high-quality data for training. The cleaned frames are then organized into batches for quick retrieval during model training, optimizing the training process and enhancing the performance of the pose estimation model. The primary objective was to provide detailed and varied set of examples for these exercises to enhance the accuracy and reliability of pose estimation models when applied to physical exercises. We have uploaded this dataset to Google Drive to facilitate direct access from Google Colab. Within the Exercise category, selective video data was organized into folders, each containing more than 100 action videos of one exercise type. All five exercise video folders—Push Ups, Jumping Jacks, Pull Ups, Body Weight Squats, and Bench Press were consolidated into a single folder named "video dataset." This aims to support the goal of detecting and analysing human activity poses across a range of upper body strength training exercises.

For the data transformation process is important to prepare the data for analysis and modelling. The transformation process includes several key steps. First, video frames are extracted from the raw video data, breaking each video into individual images. These frames are then converted into RGB format using the "**format_frame_landmarks**" function, ensuring compatibility with the BlazePose model. The model processes these RGB frames to detect body landmarks and extract their coordinates. For detecting the pose MediaPipe package "**MediaPipe PoseLandmarker**" is installed and utilized. To extract the body coordinates from each frame the BlazePose model provided by MediaPipe is employed. BlazePose is an pose estimation architecture designed to capture key movements and posture, that identifies 33 key landmarks on the human body, making it highly effective for tasks like exercise analysis and motion tracking. Moving forward, the pre-trained model pose_landmarker.task is obtained. This model, part of the MediaPipe system, uses BlazePose technology to precisely track and identify body landmarks. The extracted coordinates are arranged into structured format designed for the specific machine learning model to be used. (Sunney *et al.*, 2023)
The human pose estimation model from Figure 2. Which has the keypoints of the human body that are identified and labelled with the specific numbers corresponds to different body parts. The model has 33 keypoints starts with index 0 such as nose, eyes, ears, shoulders, elbows, wrists, hips, knees, ankles, and fingers. The connections between these points are visualized with lines, showing the relationship between the joints, essentially forming a skeletal outline of the human figure. This type of model is typically used in computer vision tasks to detect and analyze human poses in images or videos, aiding in applications like motion tracking, human-computer interaction, and sports analysis.

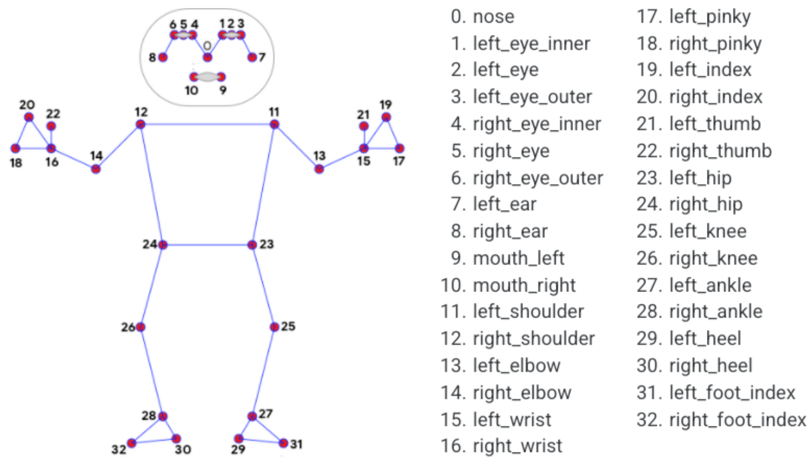

Figure .2 BlazePose Model - Topology as referenced in(Bazarevsky *et al.*, 2020)

The extracted features from the BlazePose model which converts complex coordinates data into meaningful numeric values. These features are used by various machine learning models to analyze and interpret the human physical activity or human action behaviour. Once the landmark key points are transformed into these features, those are stored as a data frame. To handle varying feature scales, Min-Max scaling is used to normalize the features between 0 and 1. This scaling ensures that no single feature disproportionately influences the model's performance. After scaling, labels are created to categorize different types of human actions. The data is then separated into dependent and independent variables and divided into training and testing sets, with 80% used for training and 20% for testing.

In data modelling phase, the models are trained using the training dataset. This research trains and evaluates three machine learning models which includes Gradient Boosting, Distributed Random Forest, Extreme Gradient Boosting Algorithm (XGBoost). (Min, 2022b)The Distributed Random Forest algorithm is applied which builds multiple decision trees during the training and merge the outputs to make the predictions. (Angelini and Naqvi, 2019)In this paper, Random Forest has provided a competitive accuracy that justified its inclusion in the comparison of methods. So, it shows that Random Forest is versatile and works well in pose estimation data which involves various features. It generates diverse trees by using random subsets of both features and training data, which helps to mitigate noise in large and complex datasets. Extreme Gradient Boosting algorithm (XGBoost) is a supervised classification technique. In a pose based human recognition research the author (Angelini and Naqvi, 2019)has utilized XGBoost algorithm that uses an ensemble of Regression Trees (CART) {R1(xi, yi)…..Rk(xi, yi)} and C Classification where Xi and yi are the training and the class label respectively. The prediction scores are summed up to get the final score. Then the final score evaluated through C additive functions, as shown in Equation:

$$\hat{y_\iota} = \sum_{k=1}^{C} f_k(x_i)_r f_k \in F$$

The final score is optimized through an objective function that combines a loss function, such as mean squared error for regression or log loss for classification, with a regularization term to prevent overfitting. This regularization includes L1 (lasso) and L2 (ridge) penalties, which control the model's complexity by penalizing large weights and encouraging sparsity. (De Paiva, Batista and Guimaraes Ramos, 2020) In the final step, Evaluation and Results, all models were evaluated based on accuracy, precision, recall and time complexity. Models are applied to the test data and the model performance was assessed using confusion matrix and report. Based on the experimentation, XGBoost classifier was determined to be the best performing model by giving highest accuracy compared to the other 2 models. And this shows that XGBoost is the suitable model to detect the human exercise pose effectively.

4 Design Specification

The purpose of this research is to enhance the accuracy and effectiveness of physical activity pose detection using machine learning techniques with classification algorithms. The proposed framework comprises several essential modules which are acquiring data and preprocessing, and feature is extracted using classification algorithms, BlazePose model, and exercise detection and evaluation. The video data capture and preprocessing module is important to analyse and observe the person's physical actions closely and preprocess them for feature extraction which includes frame selection and noise reduction as the video dataset has large number of data and stored in the cloud where the videos or images are getting compressed when dealing with the large dataset or when the bandwidth is limited. Although, the prepared data is ready to input into models like BlazePose in order to detect and extract key features related to human poses. This is the essential process to observe person's movements which helps in various domain healthcare, fitness, sports etc.

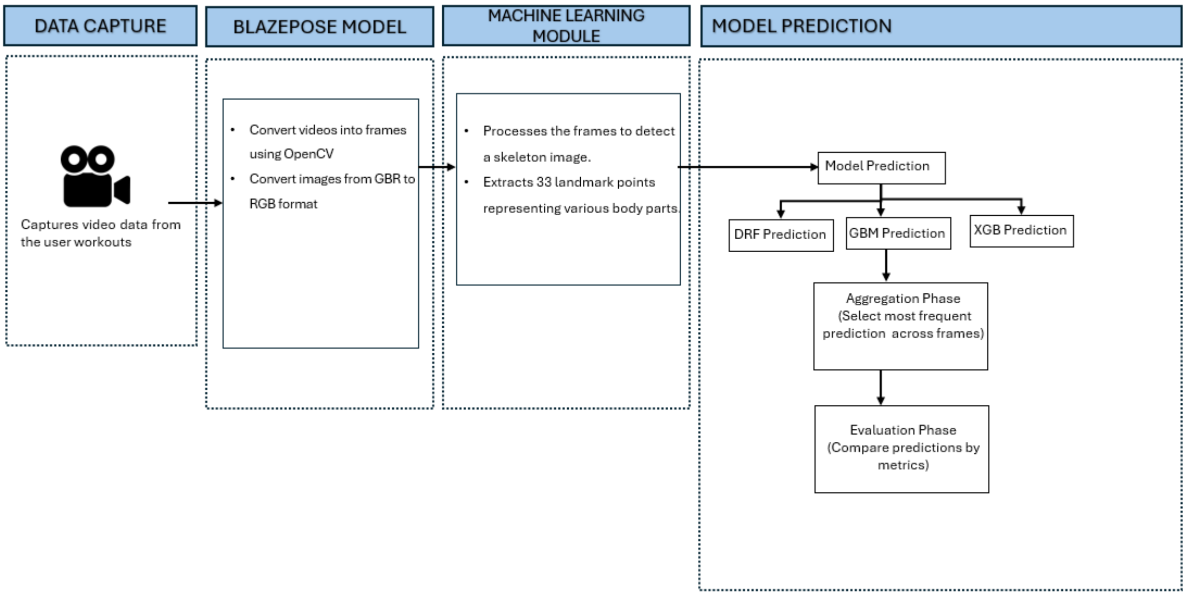

Figure.3 System Architecture

MediaPipe is used to detect an track body landmarks and movements. In this process each frame from the video data are converted from BGR to RGB format and processed to BlazePose which detects 33 key points representing a person's physical activity pose. These key points are extracted from every generated image in the video frames. Once the key points are converted into features by X and Y coordinates relative to the image dimensions, these features are then calculated by computing Euclidean distances between pairs of key points. These distances provide valuable information about the spatial relationships between different body parts. The resulting distances, along with other positional data, are combined into pose vectors. These pose vectors are then used for physical activity recognition, as well as for evaluating the model's performance in accurately recognizing and classifying these activities. These features allow the models to analyse and recognize various poses with greater accuracy.

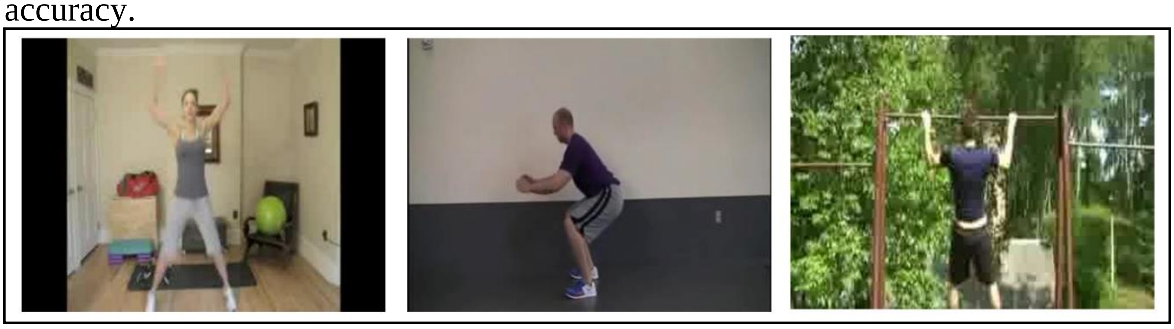

Figure.4 Original Images

In the classification module, three machine learning models are employed to analyze the strength training physical exercise pose Each of these models is used independently to predict the exercise pose and then models are combined, and the final prediction is determined based on the most frequent prediction among the three.

The performance of this ensemble prediction is then evaluated to assess its accuracy and reliability.

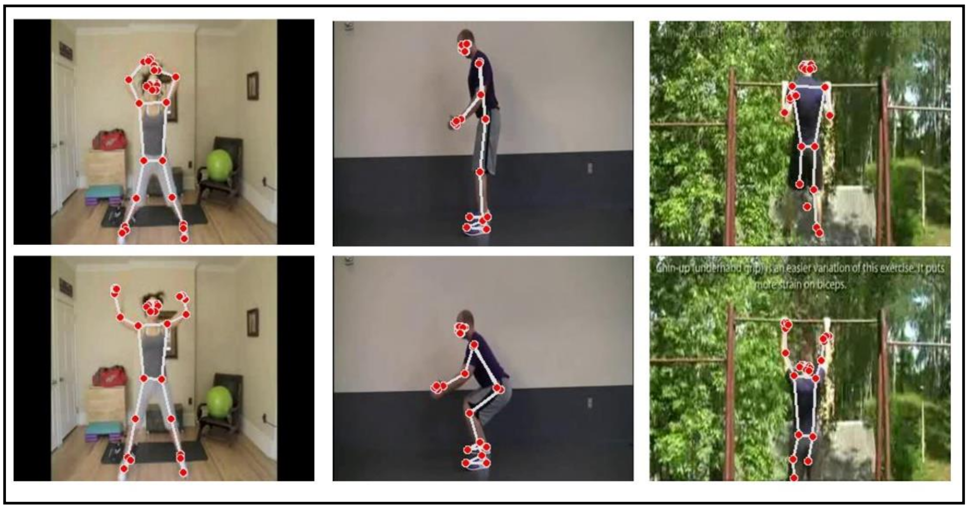

Figure.5 BlazePose Landmarks

5 Implementation

Pose-Based physical activity recognition detection was developed using the BlazePose model from the MediaPipe framework for pose estimation, and the XGBoost classifier for activity classification. This research was executed in Google Colaboratory a hosted Jupyter Notebook was utilized in this implementation of data preprocessing, data transformation, model implementation and model evaluation. The essential libraries used in this project are os, MediaPipe, NumPy, Pandas, h2o, matplotlib, and seaborn. The model is trained using the UCF101 dataset, which is a widely recognized dataset for action recognition consists of 13,320 videos covering 101 different actions performed by people in various scenarios. For this particular research, five sub-categories are taken specifically focused on strength training exercises were specifically selected for model training. Using Mediapipe's Blazepose pose estimation model preproccsed image frames are generated as 33 keypoints where each landmark provides X and Y coordinates of the body posture, The pose landmarks are extracted using **'pose_landmark()'** function and flattened the 3D landmark data into a 1D array, storing it in a CSV file to use in the machine learning model. The three machine learning models are applied and evaluated based on accuracy, precision, recall, time complexity, and model complexity. Among them, Gradient Boosting Machines (GBM) and XGBoost and (Agrawal, Shah and Sharma, 2020)Random Forest classifier performed the best for physical exercise pose detection. The final evaluation involved analyzing the most frequent predictions across video frames to determine the most accurate models. The implementation effectively combined pose estimation with exercise activity classification providing a thorough evaluation of performance while considering both model complexity and temporal dynamics of the video data.

6 Results and Evaluation

This section provides the detailed explanation of the below experiments that were conducted. The primary goal of developing the framework to predict the exercise pose using pose estimation model combined with machine learning algorithm. The performances of different machine learning algorithms were compared and analysed to determine different aspects of evaluating the model performance in classifying and predicting exercise poses.

6.1 Experiment 1

The experiment was conducted to evaluate the model performance on a separate dataset that not used during the training process. The test data is the holdout dataset which is used to assess the model's ability to generalize to new unseen data during test evaluation. Three machine learning models such as XgBoost, Random Forest, Gradient Boosting machine are considered in this experiment. The evaluation involves by calculating key performance metrics such as accuracy, precision, recall, and F1-score which includes the functions for preprocessing labels, generating a confusion matrix heatmap, and producing a detailed classification report. By comparing the model's predictions (y_pred) with the actual labels (y_true), this process determines how well the models generalize to new data, providing valuable insights into their strengths and weaknesses across different classes. Such a comprehensive assessment is crucial for determining the models' readiness for deployment.

6.2 Experiment 2

In the second experiment, Test experiment involves by evaluating the model’s performance on the individual frames of a video. In various scenarios like video analysis, each frame is treated as a separate image, and the model’s ability to make accurate predictions on these frames are assessed. This is important in applications when frame-by-frame analysis is required like tracking and pose detection during physical exercises.

This experiment aimed to test whether the classification models which have performed well in static image-based evaluation could maintain their accuracy and reliability when applied to dynamic video data. The results, as shown in Table 2 indicates that the XgBoost classifier successfully generalized video frames, by maintaining high accuracy and robustness. Additionally, the model’s ability to detect and correctly classify poses across sequential frames demonstrates its potential for real-time applications. The confusion matrix for this evaluation is shown in Figure 7, highlights the classifier's strong performance across different poses.

Classifier	Accuracy	Precision	Recall	F1
DRF	91%	91%	91%	91%
GBM	92%	92%	92%	92%
XgBoost	94%	94%	94%	94%

Table 1: Results from Experiment 2

6.3 Experiment 3

In the last experiment, The evaluation of the model's performance on entire video sequences rather than individual frames. Here the model's ability to process and analyse temporal information across multiple frames are tested. As the user performs some workout the video sequence of frames is essential. The accuracy of the model in understanding and analyzing the video as a complete sequence was thoroughly evaluated.

Classifier	Accuracy	Precision	Recall	F1
DRF	95%	95%	95%	95%
GBM	93%	93%	93%	93%
XgBoost	96%	96%	96%	96%

Table 3: Results from Experiment 3

The below figure 6 compares the exercise predictions made by three machine learning models using the video labelled as "weighted squats" While mode models consistently predict the correct exercise, a discrepancy is noted where the GBM model incorrectly predicts "jumping jacks" for one of the videos, whereas the other models correctly identify it as "weighted squats." This comparison highlights the performance differences among the models in classifying the exercise types.

1 to 5 of 5 entries Filter

index	exercise	vid name	drf_prediction	gbm_prediction	xgb_prediction
0	weighted squats	v_BodyWeightSquats_g18_c04.avi	weighted squats	weighted squats	weighted squats
223	weighted squats	v_BodyWeightSquats_g17_c01.avi	weighted squats	weighted squats	weighted squats
403	weighted squats	v_BodyWeightSquats_g15_c02.avi	weighted squats	weighted squats	weighted squats
525	weighted squats	v_BodyWeightSquats_g24_c02.avi	weighted squats	weighted squats	weighted squats
746	weighted squats	v_BodyWeightSquats_g22_c01.avi	weighted squats	jumping jacks	weighted squats

Show 25 per page

Figure 6: Comparison of Exercise Predictions

7 Discussion

This section describes the detailed discussion of the findings which is acquired from all the experiments conducted during this study. For this study UCF 101 dataset is utilized to detect and predict the exercise pose in which individual done during the strength training workout. Mediapipe's BlazePose model was utilized to extract key point features, which were then converted into H2O frames to analyze the relationships between key points and body angles. This framework consists of three machine learning models where it predicts the exercise pose with which it counts on each rep and detect the pose at each rep and finally it is predicted by deciding the most frequent pose across the frames. Meanwhile, the time taken for this prediction is tracked and calculated for each prediction in seconds. All the above experiments are conducted using Google Colab with 16 GB RAM. Based on the results of the study it shows that XgBoost classifier is the best fit for the physical exercise activity pose detection model with the accuracy of 94% test evaluation and the accuracy of 96% in the frame-by frame evaluation. Using the low-quality videos XgBoost classifier gives best accuracy than the other algorithms.

The first experiment was conducted to test the data using XgBoost, Random Forest, and Gradient Boosting models. The varying strengths and weaknesses of these models highlight their suitability for different deployment scenarios. The evaluation was performed using both (Amrutha, Prabu and Paulose, 2021) static and dynamic input, assessing the models frame by frame. The second experiment analyzed which model was the best fit for physical activity pose detection using the BlazePose framework. The third experiment evaluated the pose detection model based on its predictions. This experiment aimed to assess the model's ability to accurately analyze and predict poses in dynamic scenarios, confirming its effectiveness for real-time applications. Although the framework's results are generally good, the model occasionally gets confused with certain exercise poses during some repetitions. Improving this aspect could enhance the model's overall performance and accuracy.

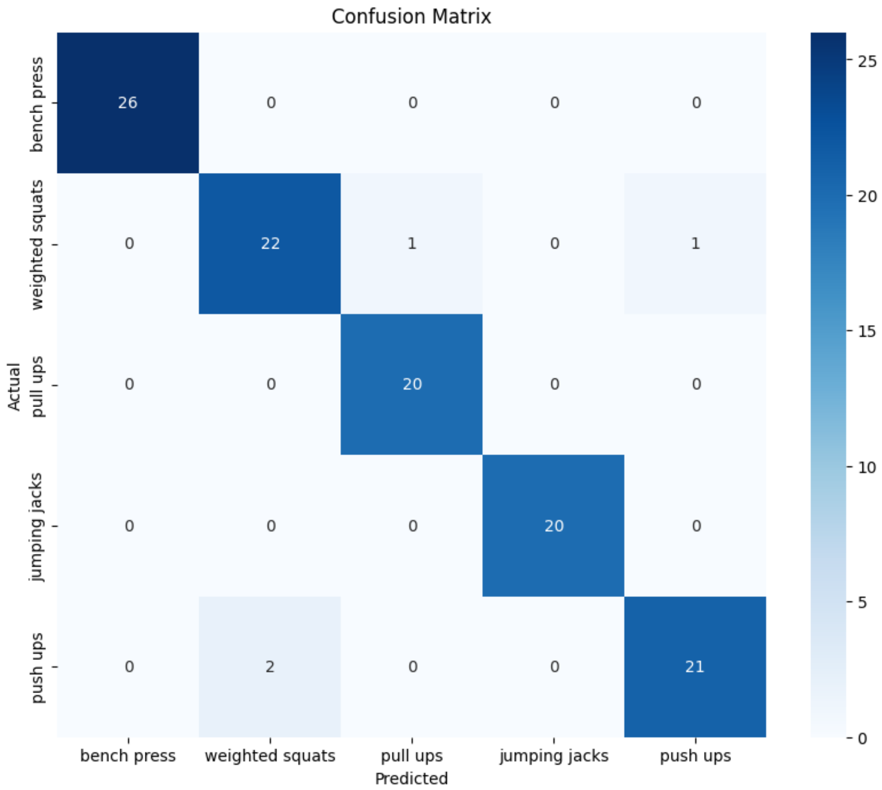

Figure .7 Confusion Matrix for human physical activity pose detection model

8 Conclusion and Future Work

The aim of this research was to predict the exercise pose by combining BlazePose model with the machine learning approach. A comparison of the accuracy of all the three machine learning models was performed in this research. Machine learning models such as Extreme Gradient Boosting model (XGBoost), Distributed Random Forest, Gradient Boosting Machine (GBM) gave an accuracy of 95%, 93%, 96% respectively. For this research dataset used from UCF 101 Human Action Recognition category utilized five selected exercise workout videos in .avi format: Pushups, Pullups, Body weight squats, Jumping Jacks, Bench press. Based on the research findings, it was concluded that physical activity of strength weight exercise pose detection can be effectively accomplished by integrating classification machine learning models with BlazePose estimation framework.

In this study, MediaPipe and OpenCV are utilized to detect the pose landmarks from images or videos. OpenCV handled the input images and videos and converted the frames to RGB format for compatibility with MediaPipe. To implement and run the MediaPipe model on a system with 8GB of RAM and an Intel i5 multi-core processor, The PoseLandmarker model from MediaPipe analyzed the frames to extract the pose landmarks, which is assigned with x,y and z coordinates. The format_frame_landmarks function processed each frame or image to collect these coordinates, while the extract_pose_landmarks function managed the file input and organized the collected landmark data into Pandas DataFrame for further analysis. This approach seamlessly integrates computer vision and data analysis tools to enhance the accuracy and utility of pose detection and interpretation. The framework uses MediaPipe for detecting pose landmarks in images or videos, which allows it to track human body movements and uses geometric features for training and classification. This helps to detect the human activity poses of strength training workouts in complex environments. This research results demonstrate that the framework excels in detecting and accurately classifying specific types of exercise poses.

For the future work, the framework can be deployed using cloud platforms such as AWS, Google Cloud, or Azure for large sensitive data processing, followed by data governance and ethical considerations. Creating RESTful APIs with eh tools like Flask for real-time predictions can enhance user satisfaction. Monitoring the activity performed by the users can track metrics such as latency and error rates to maintain the accuracy of the performance. (Wang, 2024a) A real-time dashboard can be created to display performance logs within the service package or the user's account, enabling self-analysis. Incorporating user feedback and monitoring costs are the key steps to optimize the model's effectiveness and efficiency over time.

References

Agrawal, Y., Shah, Y. and Sharma, A. (2020) 'Implementation of machine learning technique for identification of yoga poses', *Proceedings - 2020 IEEE 9th International Conference on Communication Systems and Network Technologies, CSNT 2020*, pp. 40–43. Available at: https://doi.org/10.1109/CSNT48778.2020.9115758.

Amrutha, K., Prabu, P. and Paulose, J. (2021) 'Human Body Pose Estimation and Applications', *3rd IEEE International Virtual Conference on Innovations in Power and Advanced Computing Technologies, i-PACT 2021* [Preprint]. Available at: https://doi.org/10.1109/I-PACT52855.2021.9696513.

Angelini, F. and Naqvi, S.M. (2019) 'Joint RGB-Pose Based Human Action Recognition for Anomaly Detection Applications', *FUSION 2019 - 22nd International Conference on Information Fusion* [Preprint]. Available at: https://doi.org/10.23919/FUSION43075.2019.9011277.

Bazarevsky, V. et al. (2020) 'BlazePose: On-device Real-time Body Pose tracking'. Available at: https://arxiv.org/abs/2006.10204v1 (Accessed: 11 August 2024).

Bukht, T.F.N. and Jalal, A. (2024) 'A Robust Model of Human Activity Recognition using Independent Component Analysis and XGBoost', *2024 5th International Conference on*

Advancements in Computational Sciences, ICACS 2024 [Preprint]. Available at: https://doi.org/10.1109/ICACS60934.2024.10473238.

Irfan, I. and Abdul Muthalib, M. (2022) ‘Implementation of Human Pose Estimation Using Angle Calculation Logic on The Elder of The Hands as A Fitness Repetition’, *International Journal of Engineering, Science and Information Technology,* 2(4), pp. 101–110. Available at: https://doi.org/10.52088/IJESTY.V2I4.346.

Krishnan, H. et al. (2022) ‘Pose Estimation of Yoga Poses using ML Techniques’, *INDICON 2022 - 2022 IEEE 19th India Council International Conference* [Preprint]. Available at: https://doi.org/10.1109/INDICON56171.2022.10040162.

Min, Z. (2022b) ‘Human Body Pose Intelligent Estimation Based on BlazePose’, *2022 IEEE International Conference on Electrical Engineering, Big Data and Algorithms, EEBDA 2022,* pp. 150–153. Available at: https://doi.org/10.1109/EEBDA53927.2022.9745022.

Mroz, S. et al. (2021) ‘Comparing the Quality of Human Pose Estimation with BlazePose or OpenPose’, BioSMART 2021 - *Proceedings: 4th International Conference on Bio-Engineering for Smart Technologies* [Preprint]. Available at: https://doi.org/10.1109/BIOSMART54244.2021.9677850.

Nunes, U.M., Faria, D.R. and Peixoto, P. (2017) ‘A human activity recognition framework using max-min features and key poses with differential evolution random forests classifier’, *Pattern Recognition Letters, 99,* pp. 21–31. Available at: https://doi.org/10.1016/J.PATREC.2017.05.004.

De Paiva, P.V.V., Batista, M.R. and Guimaraes Ramos, J.J. (2020) ‘Estimating Human Body Orientation using Skeletons and Extreme Gradient Boosting’, *2020 Latin American Robotics Symposium, 2020 Brazilian Symposium on Robotics and 2020 Workshop on Robotics in Education, LARS-SBR-WRE 2020* [Preprint]. Available at: https://doi.org/10.1109/LARS/SBR/WRE51543.2020.9307079.

Sajan, S. and Anilkumar, A. (no date) *2 nd International Conference on IoT Based Control Networks and Intelligent Systems (ICICNIS 2021),* ‘Pose Estimated Yoga Monitoring System’. Available at: https://ssrn.com/abstract=3882498 (Accessed: 11 August 2024).

Setiyadi, S. et al. (2022) ‘Human Activity Detection Employing Full-Type 2D Blazepose Estimation with LSTM’, APWiMob 2022 - *Proceedings: 2022 IEEE Asia Pacific Conference on Wireless and Mobile* [Preprint]. Available at: https://doi.org/10.1109/APWIMOB56856.2022.10014270.

Singh, D., Panthri, S. and Venkateshwari, P. (2022) ‘Human Body Parts Measurement using Human Pose Estimation’, *Proceedings of the 2022 9th International Conference on Computing for Sustainable Global Development, INDIACom 2022,* pp. 288–292. Available at: https://doi.org/10.23919/INDIACOM54597.2022.9763292.

Song, Z. and Chen, Z. (2024) ‘Sports Action Detection and Counting Algorithm Based on Pose Estimation and Its Application in Physical Education Teaching’, *Informatica, 48(10),* pp. 35–50. Available at: https://doi.org/10.31449/INF.V48I10.5918.

Sunney, J. et al. (2023) ‘A Real-Time Machine Learning Framework for Smart Home-based Yoga Teaching System’, *Proceedings - 2023 7th International Conference on Machine Vision and Information Technology, CMVIT 2023,* pp. 107–114. Available at: https://doi.org/10.1109/CMVIT57620.2023.00029.

Wang, T. (2024b) ‘Intelligent Long Jump Evaluation System Integrating BlazePose Human Pose Assessment Algorithm in Higher Education Sports Teaching’, *Systems and Soft Computing*, p. 200130. Available at: https://doi.org/10.1016/J.SASC.2024.200130.

Zhang, S. et al. (2022) ‘Human deep squat detection method based on MediaPipe combined with Yolov5 network’, *Chinese Control Conference, CCC*, 2022-July, pp. 6404–6409. Available at: https://doi.org/10.23919/CCC55666.2022.9902631.