IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 5 June 2024, accepted 23 June 2024, date of publication 15 July 2024, date of current version 29 July 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3428309

RESEARCH ARTICLE

Time-Series Data to Refined Insights: A Feature Engineering-Driven Approach to Gym Exercise Recognition

AFZAAL HUSSAIN[1], MUHAMMAD ADEEL ZAHID[2], USAMA AHMED[3], SHAHBAZ NAZEER[1], KASHIF ZAFAR[4], AND ABDUL RAUF BAIG[5], (Member, IEEE)

[1]Department of Information Technology, Government College University Faisalabad, Faisalabad 38040, Pakistan
[2]Department of Computer Science, Government College University Faisalabad, Faisalabad 38040, Pakistan
[3]Department of Software Engineering, Government College University Faisalabad, Faisalabad 38040, Pakistan
[4]Department of Computer Science, National University of Computer and Emerging Sciences, Islamabad 54770, Pakistan
[5]Department of Information Systems, College of Computer and Information Sciences, Imam Mohammad Ibn Saud Islamic University (IMSIU), Riyadh 11564, Saudi Arabia

Corresponding author: Usama Ahmed (usamaahmed@gcuf.edu.pk)

ABSTRACT Machine learning-based sports activity recognition has captured a lot of interest in recent years. Automatic activity recognition not only reduces cost and time but is very helpful in analyzing health-sensitive data acquired using smart wearable technology. Gym activity recognition by incorporating smart wearable technology comes within the scope of this topic. This paper present a system for classifying gym activities using feature engineering techniques applied to time series data. The collected time series data consists of an athlete's body movement using an internal 3-axis accelerometer build into the zephyr bio-harness 3 device. The data were gathered by implementing a six-week fitness routine trying to target six muscle groups, preceded by one day of rest and recovery each week. The raw time-series data of the accelerometer is transformed to extract new features from it for identifying gym activities. The feature engineering techniques applied in this research are not limited to gym activity recognition but can be extended to any domain involving time-series data. The collected data was just three features, which are the reading of the tri-axial accelerometer signal as vertical, lateral, and sagittal axes. In order to formulate new features, basic concepts of statistics and mathematics were applied to the data. furthermore, we trained six GridSearchCV-based classifiers on the extracted features and tested their performance in four different types of experiments.

INDEX TERMS Internet-of-Things (IoT), feature engineering, gym activity recognition, machine learning, smart workout.

I. INTRODUCTION

The human beings are capable of completing activities on multiple levels, including mental, emotional, and physical. Any action that improves or maintains a practitioner's health and fitness is considered a physical exercise. The exercises are divided into three categories: flexibility exercises, aerobic exercises, and anaerobic workouts. A gym is a technologically advanced indoor sports facility that is outfitted with a variety of apparatus that is used by bodybuilders, strength trainers, weight lifters, and practitioners. In recent years, there has been a lot of interest in human activity recognition (HAR). Human activity detection is commonly performed based on smartphone sensor data i.e. accelerometer and gyroscope. These human activity recognition techniques are applicable in the health monitoring of athletes, remote patient monitoring, elderly care, telemedicine services, and smart environments [2], [3]. This paper considers a very specific case of HAR in health and fitness domain and focuses on gym exercise recognition (GER). Unlike general HAR, gym exercise recognition requires highly precise monitoring to capture and correlate intricate details of complex body movements that may be common to multiple exercises.

The associate editor coordinating the review of this manuscript and approving it for publication was Daniel Augusto Ribeiro Chaves.

In machine learning, a feature plays an important role between data and models. Machine learning-based prediction models use scientific models on data to derive insights by taking features as input. The process of extracting new transformed features from raw data that are applicable to the machine learning model is known as feature engineering. Feature engineering is a crucial process because the right features can ease the difficulty of machine learning models, and enable them to generate results of higher quality. Well-feature engineering and data cleaning consumes the majority of the time in a machine-learning model. Yet, feature engineering is rarely discussed on its own [4].

In machine learning, feature engineering is the core activity that is performed to construct suitable features to make data ready for classification purposes for better prediction. Feature engineering uses transformation methods like arithmetic, and statistics on a set of features to generate new features. Such transformations in feature engineering help scale a feature that is easier to learn. In other words, feature engineering transforms the feature space and improves the predictive performance of a given dataset [5]. Human Activity recognition includes data segmentation, feature extraction, feature selection, and classifier training phases. Feature extraction is the major phase among all the phases mentioned above as it affects the accuracy of the classifier. Feature extraction from the smart sensors data is a challenging task due to the variable orientation of smartphones, placement and subject [6].

This research concentrates on gym exercise recognition by applying feature-engineering techniques to the accelerometer data. Feature engineering is the process of transforming time series data into features that better represent the underlying problem to the predictive models, thereby improving their performance. The ultimate goal is to make gym trainers' lives as easy as possible by implementing automation using current breakthroughs in the Internet of Things (IoT) and associated technologies. The great majority of human activity research focuses on only a few daily tasks. These activities are considerably different from the exercises performed in a gym using gym equipment. As a result, we created our own dataset.

The contribution of this paper is three folds and discussed as follows.

- We utilized smart wearable devices equipped with an accelerometer to capture time-series data related to exercises performed in a gym setting in real-time.
- Despite having only limited set of features extracted from the raw accelerometer data (such as acceleration in three axes: vertical, lateral, and sagittal), we transformed this raw time-series data into a set of 106 distinct features that can be used for analysis and modeling. This rich set of features extracted from limited time-series sensor data has enhanced the accuracy and effectiveness of our Gym Exercise Recognition (GER) system.
- Traditional classification methods such as Random Forest, Decision Tree, and Support Vector Machine cannot be directly applied to time series data due to its sequential and temporal nature. Our research has tried to establish a benchmark for adapting traditional classification algorithms to time series data, thereby expanding the range of methods available for analysis and potentially improving predictive performance.

Further to this introduction, remaining paper has the following structure. The Background section presents a review of the literature on gym exercise recognition research. The Material and Methods section presents details of the collected dataset and the transformation process using feature engineering. Our experimental design decisions, which were made based on the various options available for recognizing gym exercises in our collected dataset, are presented in Experimentation section. The Results section provides the experimental validation of the gym exercise recognition system on the transformed data followed by a Conclusion section.

II. BACKGROUND

The goal of this research is to try and recognize gym exercises using data from wearable sensors and feature engineering on time series data. The following research on gym activity identification and feature transformation is important to discuss in order to put our findings in perspective.

A. FEATURE ENGINEERING OF ACCELEROMETER DATA

Data segmentation, feature extraction, feature selection, and classifier training phases are all involved in human activity recognition. Feature extraction is the major phase among all the phases mentioned above as it affects the accuracy of the classifier. The related research is reported in [3], [4], [5], [6], [7], [8], and [9]. A feature engineering framework named Learning Feature Engineering (LFE) is proposed in [5] that can build new features based on learning patterns between feature transformations, feature characteristics, and class distributions. The feature representation module in LFE is Quantile Sketch Array (QSA), which can produce fixed-size arrays, preserving essential characteristics of features at a low computational cost for a variety of classification problems. In 2017, Quiroz et al. [3] compared the accuracy achieved with features extracted for activity recognition on a feature-engineered dataset. They analyzed accuracy over gravity signals, Jerk signals, and angular velocity. Their feature engineering work is beneficial for both data extracted from smartphones and human activity recognition. The research presented in [6] provides a feature extraction and reduction model based on EPS and LDA for the smart sensor dataset. The prime target was to extract quality features from raw sensing data. The suggested model is useful for noise reduction and feature extraction from accelerometer and gyroscope data. For the purpose of evaluating the model's performance on publicly accessible datasets, a supervised classification model is included. The model outperformed previous feature extraction techniques in terms of cost and performance.

B. GYM EXERCISE RECOGNITION

Recofit [10], is a wearable sensor-based system that presented a system for recognizing and counting the repetitions of gym exercises using data from a wristband with a 3-axis accelerometer and gyroscope. A free weight monitoring system (FEMO) is proposed in [11] the system can recognize gym activity and count the number of repetitions. IMU-based solution [12] acquired athlete's movements during exercises using a Bluetooth-based device and performed classification, segmentation on groups of exercises, and counted exercises repetitions. A smart watch-based system MiLift [13] tracks gym exercises and categorizes gym exercisers into different classes walking, running, exercising, and resting.

The research presented in [14] uses the built-in accelerometers of smartphones to collect data on physical activity with the goal of determining the beginning, end, and length of repetition from continuous acceleration. Both a constrained and an unconstrained setting is used in the experiments. While the unconstrained workouts don't call for specific gym equipment, the activities in the limited setting were carried out in a gym using weight machines. These workouts don't require any special equipment and may be done anywhere, including outside, in a park, and in a gym. The research reported in [15], addresses the issue of rating the effectiveness of weightlifting exercises and offering feedback on them. In order to define the quality of execution, they investigated 3 aspects of gym exercise recognition. These aspects are correct execution, execution mistakes, and user feedback. The data collection process consists of five variations of the exercise. In each variation, one set of 10 repetitions is performed. Among these variations, one of the variations is related to the specified execution while the other is mistakes. Four inertial measurement units (IMU) sensors consisting of three-axis acceleration, gyroscope, and magnetometer data are placed in the athlete's glove, armband, lumbar belt, and dumbbell.

A smart wearable-based exercise and intensity recognition system is proposed in [16]. They attempt to recognize the intensity of six upper body strength training exercises. A network of five body-mounted accelerometers is used to record the acceleration and a smartphone acts as a hub for data reception and processing. The sensor's placement varies by chest, upper arms, and wrists. The automated tracking and analyzing of weight training exercises work reported by [17] aiming for accurate and fast-tracking of exercises. They targeted seven dumbbell exercises with ten participants. A prototype hardware consisting of accelerometers and gyroscopes is used to perform exercises using a dumbbell. Automatic gym exercise recognition and quantification have been attempted in [12]. The study examines data from wrist-worn wearable IMUs and makes an effort to categorize and count exercises with a focus on nine gym routines.

The research that was just described has a lot to do with weight training in a gym setting. A large body of literature on the detection and analysis of sports-related activities and fitness activities is present in [18], [19], [20], [21], [22], [23], [24], [25], [26], and [27]. These paper are based on a systematic review of the literature on machine learning for sports activity recognition using different devices. Likewise, there are several excellent surveys on the primary topic of activity recognition [28], [29], [30].

III. MATERIAL AND METHODS

The gym related data acquisition is a challenging task in human activity recognition research. The first difficulty is in finding the most appropriate participants, preparing them, and keeping them intact during the whole data collection phase. The cost of sensors and privacy concerns are some common issues. We collected gym-related data in our research reported in [31] and [32] using the commercially available device zephyr bioharness [33]. The information gathered is largely about bodybuilding and weight training methods. These exercises are regularly utilized by free-weight practitioners and have well-known descriptions. This section illustrates the entire data-collecting and transformation process in a gym setting.

A. THE DATA COLLECTION PROCESS

The smart wearable-based data collection process involves forty-four gym exercises arranged into six muscle groups. The Government College University, Faisalabad, Pakistan's [34] Ethical Review Committee approved the data collecting process. The data is collected in 2019 and the data related to individual athlete can be identified by the authors. The athletes were selected with ages between 26 to 39 years and weights between 74 to 105kg. Each athlete was given a weekly workout routine targeting all the muscle groups of the athlete's body. In the proposed weekly workout routine, three sets of all the exercises related to a single muscle group are performed each day. Six workouts are scheduled for six consecutive days of the week, with the seventh day functioning as a rest day. The data acquisition process spanned six weeks as per the workout plan provided to the athletes. The athlete performed a total of forty-four exercises in a week. The workouts were divided into six groups based on the targeted muscle group. Arms, back, chest, shoulder, legs, and core-body are the muscle groups involved. The first five muscle groups necessitate the use of gym equipment. The core-body muscle group is associated with activities that do not require gym equipment and can be performed anywhere. Core-body workouts can help you increase your stamina, flexibility, and balance. Each workout corresponds to a different day of the week and must be completed in a single session by executing three sets of each exercise. A set is a sequence of repetitions of an exercise. The length of each exercise and set is determined by the level of fitness of the athlete. Selected exercises are common and well-known.

A Zephyr Bio Harness 3 (BH3) [33] device is used to capture bodily movement data. The devices are equipped with 3-axis accelerometers that measure acceleration in all three spatial dimensions, namely vertical, lateral, and sagittal. The accelerometer on the gadget can identify the

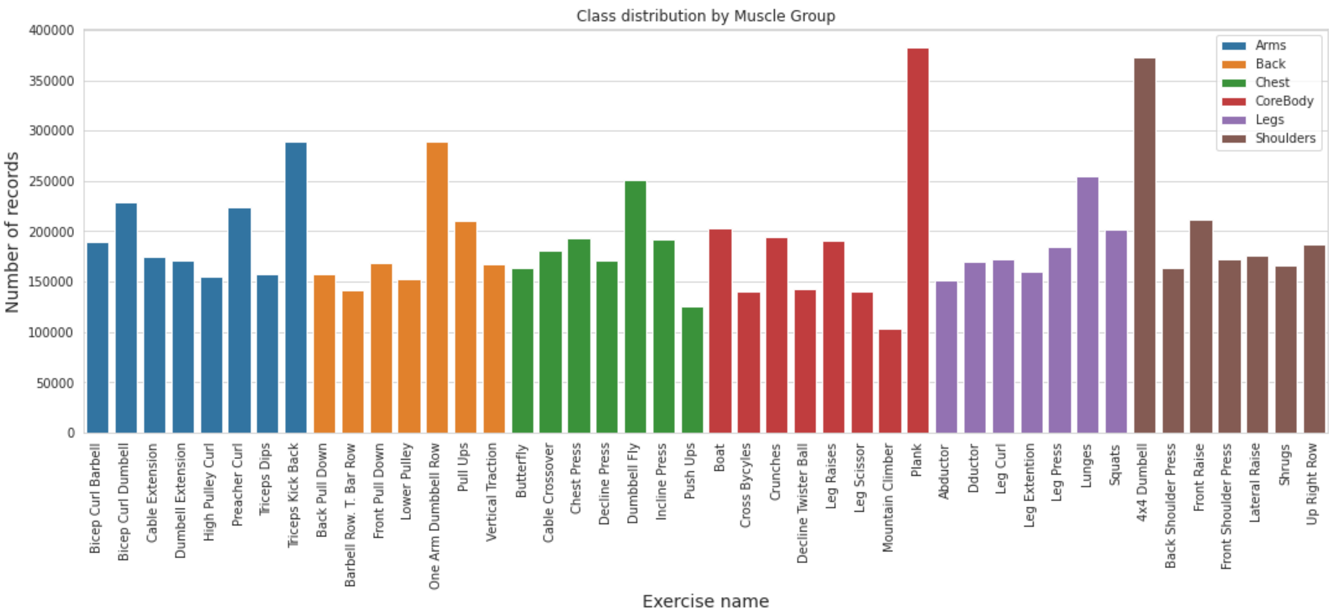

FIGURE 1. **Exercise-wise instance count based on the muscle group.**

athlete's orientation, which can be useful for recognizing gym exercises. During each workout session, the device was fastened to an athlete's chest to collect real-time body movement data. A mobile application is used to collect workout data. The information includes the athlete's name, exercise name, start and end times, set numbers, and the exercise apparatus utilized. This workout data was later appended to the raw BH3 signal data to complete the dataset.

The body posture data was recorded using the BioHarness device's accelerometer. While the workout data was captured using the mobile application. The Body posture data are accelerometer readings along the vertical, lateral, and sagittal axis at a given timestamp. The body movement data is received at 100Hz frequency, that is 100 samples per second. This data consists of 4 features, the features are *timestamp*, *vertical*, *lateral*, and *sagittal* axis. The workout data consists of mainly six attributes: start and end timestamp of an exercise, user, workout id, workout number, and exercise name. Because labeled data is required for training and testing supervised learning-based classification models, each transformed instance is tagged with the exercise completed based on workout data collected from the mobile app. The exercise name is the variable we want to predict.

The collected time series dataset comprised of 8,396,799 samples, each consist of 7 attributes, four attributes from the accelerometer and three attributes are mapped from mobile application for labelling.The class label distribution for the time series data is depicted in Figure 1 for each exercise performed as per their muscle group. As the gym exercises are continuous, so it can be observed from the figure that the time for each exercise performed varies. This won't affect the overall activity recognition as we have a significant amount of data samples.

B. DATASET DESCRIPTION

The dataset is in two categories as listed below.

1) Body Posture Data.
2) Workout Data.

For a particular exercise, let's observe how the signal values in each of the vertical, lateral, and sagittal dimensions vary with time. A subset of 1000 samples is considered for visual inspection of the signal. This sample is equivalent to 10 secs of the exercise as the frequency of data collection was 100 Hz. Figure 2 presents a visual inspection of chest press and decline press exercises. As you can notice, the signal shows periodic behavior for the decline press exercise while it has very less but different movement for the chest press exercise. The activity-wise distribution of the signal data along the vertical, lateral, and sagittal axes is shown in Figure 3 to examine if there is any evident pattern based on the range and distribution of the values. It has been shown that there is a significant amount of data overlap between activities such as bicep curl dumbbell and dumbbell extension on all the axis. Tricep dips, tricep kickbacks, and preacher curl appear to have distinctive values along the vertical, lateral, and sagittal axis.

C. DATA TRANSFORMATION

As you can see in Figure 2 from the raw accelerometer data the prediction of the gym activity depends upon only three variables vertical, lateral, and sagittal, these variables are time-series based. Conventional classification algorithms cannot be used directly for such data. Instead, we must first use feature engineering techniques to convert the raw time-series data. The process of converting raw signal data and developing new features from it is known as feature engineering. The data that has been modified can be used

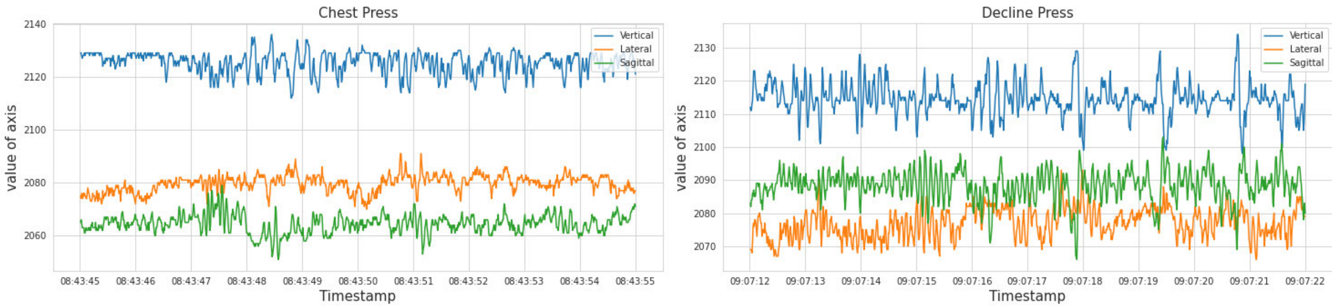

FIGURE 2. Visual inspection of acceleration data for Chest Press and Decline Press activity from chest workout.

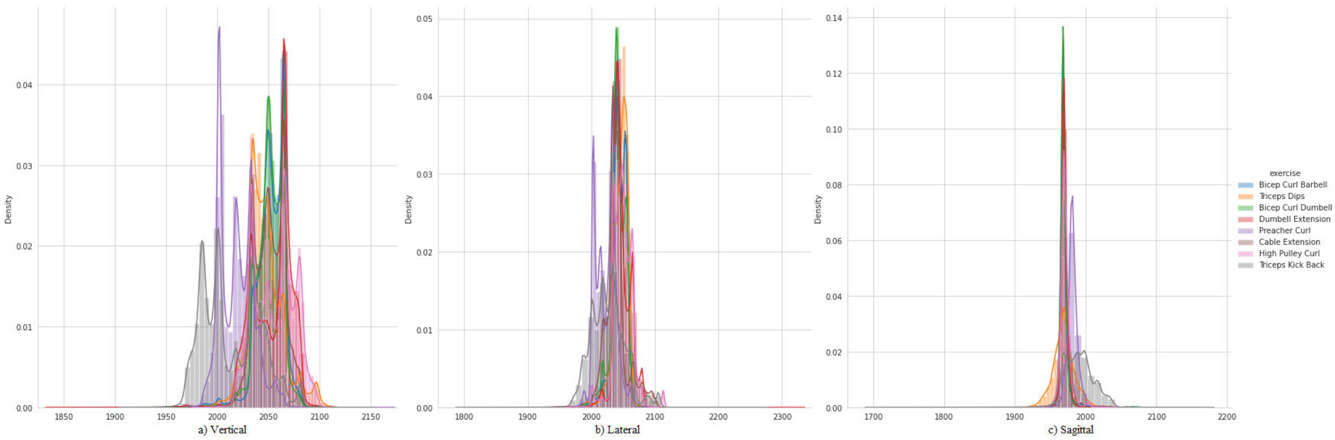

FIGURE 3. Exercise-wise distribution of axes.

by the Gym Activity Recognition module. This section presents strategies for transforming raw time-series data and extracting new features from it. The strategies presented in this article are not confined to the prediction of gym activity but can be applied to any domain requiring time-series data. The architecture of the proposed system for gym exercise recognition is presented in Figure 4. In order to transform the raw time-series data windowing technique is used as presented in Figure 5. This illustrates how the raw signal data is collected and turned into new features. This method aggregates the 400 raw samples that are contained in each of the four-second windows of data with a step size of one second or 100 samples to produce new features. The window's most frequent exercise is chosen to serve as the class label for the elements that have been modified.

D. FEATURE ENGINEERING

In this section, we are going to transform three accelerometer variables i.e vertical, lateral, and sagittal axis of our dataset to design new features. As we know, in classification we need to split data into train and test datasets, typically we can do splitting based on random sampling on whole data but in our scenario, we are going to split the dataset before transformation on *setno* attribute as it makes more sense. The attribute *setno* is important as each user performs three sets of an exercise in a day. So we have two sets as training data and one set as test data. This is just like splitting 2/3 data samples in the training set and 1/3 samples in the test set in a more sophisticated manner.

Feature engineering is applied in three stages on each four-second window based on transformed data as discussed in the previous section and shown in Figure 5. Each window contains 400 samples of the vertical, lateral, and sagittal axis which is equal to 1200 samples in a window. Each window is processed to create one record of new features. The stages of features engineering are described below in sub-sections.

1) STATISTICAL MEASURES

In statistical measure based feature engineering we applied statistical measures listed below on each window of transformed dataset and build new features. The statistical measures are listed below.

1) Mean.
2) Standard deviation.
3) Average absolute deviation.
4) Minimum value.
5) Maximum value.
6) Difference of max and min values.
7) Median.
8) Median absolute deviation.

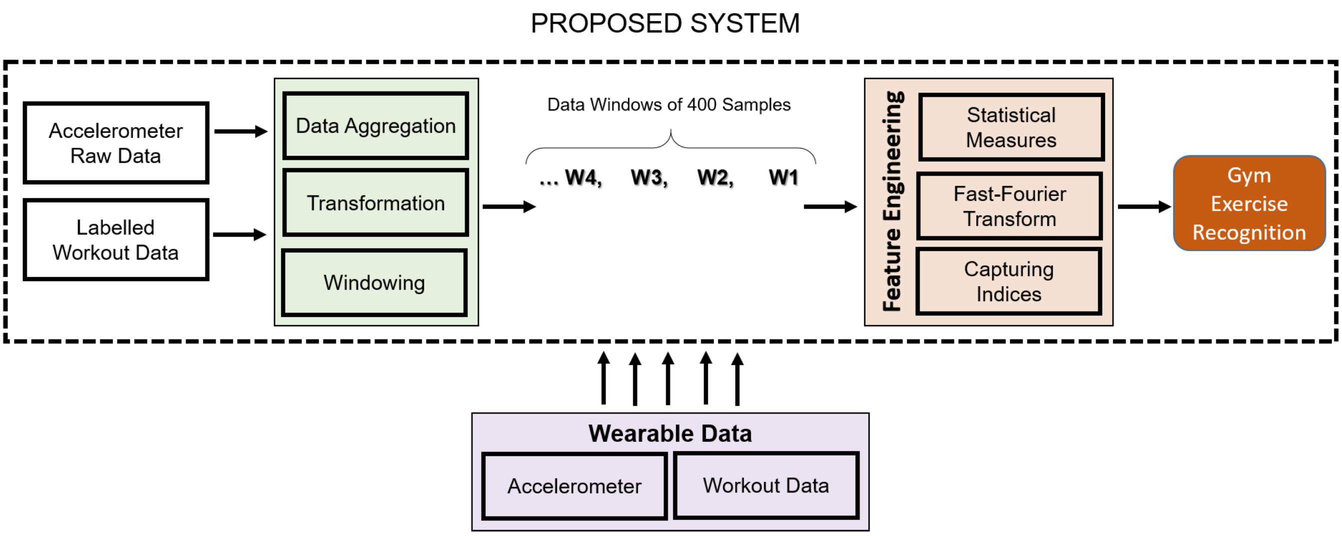

FIGURE 4. **The Proposed System for Gym exercise recognition.**

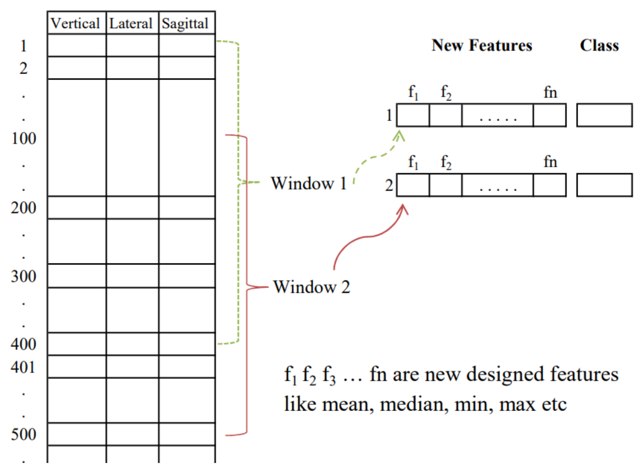

FIGURE 5. **Transformation of raw data using windowing technique.**

9) Interquartile range.
10) Number of values above mean.
11) Number of peaks.
12) Skewness.
13) Kurtosis.
14) Energy.
15) Average Resultant Acceleration (ARA).
16) Signal magnitude area (SMA).

A total of 16 statistical measures were applied to construct a new dataset comprising 44 features. The first fourteen measures were applied individually to each of the three axes (vertical, lateral, and sagittal), resulting in forty two features. The remaining two measures ARA and SMA each generated one additional feature.

The majority of these statistical measures are obvious and widely known. The function that takes a 1-D array and discovers all local maxima by straightforward comparison of surrounding values produces the peaks for each axis inside a signal. Two statistics that shed light on the distribution's form are skewness and kurtosis. The mean of the sum of squares of the data in a window along each axis is used to calculate the energy of a signal along that axis. The average of the square roots of the values in each of the three axes squared and combined together is used to get the average resultant acceleration over the window. The total absolute values of the three axes averaged over a window are used to calculate signal magnitude area.

We'll be using this new dataset in the sections that follow, gradually adding more features to it, and ultimately using it to train machine learning models.

2) FAST-FOURIER TRANSFORM

A time domain signal is converted into a frequency domain signal using the Fourier transform function. The function takes a temporal signal as input and outputs the signal's frequency representation. In the real world, every signal is a temporal signal made up of several sinusoids with various frequencies. The signal is not changed by the Fourier transform. Though some characteristics and elements of the signal may be completely studied in the frequency domain, it merely offers a fresh perspective for examining your time signal.

In the previous section we only dealt with the time domain and computed forty four features. Here, we applied FFT on the computed features and with FFT forty four additional features were computed. This resulted in a total of 88 features. This development marks Stage 3 of our feature engineering process.

3) CAPTURING INDICES

In any machine learning challenge, the major objective of the feature engineering step is to give the model as much data as you can. It learns more effectively when you give it additional knowledge. The underlying data's index values have been chosen as prospective features and are listed for each axis below.

1) Index of max value in the time domain
2) Index of min value in the time domain
3) Absolute difference between the above two indices
4) Index of max value in the frequency domain
5) Index of min value in the frequency domain
6) Absolute difference between the above two indices

These six measures were applied on the previously computed attributes, resulting in eighteen new features. Now, at the end of stage 3, we have generated 18 more features in our dataset.

E. TRANSFORMED DATA: ENGINEERED FEATURES

Based on the feature engineering process as discussed in the previous section, we derived 106 features from the initial 3 features of the raw accelerometer data. In addition to these derived features, there are four more features mapped from the mobile application. The feature-engineered dataset ultimately consisted of 83,964 samples, each comprising 110 attributes.

IV. EXPERIMENTATION

It's time to test how effectively these recently crafted features predict gym exercises. The transformed data is used for experiments related to gym exercise recognition. This section describes the classification algorithms and the methodology of experiments on gym exercise recognition.

A. METHODOLOGY OF THE EXPERIMENTS

Gym exercise recognition entails locating a specific activity that is part of a set of exercises. In the transformed data set, exercise represents a class, this means we have a total of 44 classes. The *setno* attribute is used to partition the modified training and test data so that each athlete's data is represented equally in both the training and test sets. According to the athlete and muscle group stated below, we have set up four types of experiments.

1) Experiment 1: Muscle group dependent and Athlete independent
2) Experiment 2: Muscle group and Athlete dependent
3) Experiment 3: Muscle group and Athlete independent
4) Experiment 4: Muscle group independent and Athlete dependent

In our first experiment, the transformed data for each muscle group consists of exercise data from all the athletes including exercises related to that muscle group. The muscle group-dependent exercise models contain data from all the athletes and we have a total of 6 models one for each muscle group. In the second trial, each athlete receives his or her own data that includes exercises specifically for each of the six muscle groups. As a result, we have data on each athlete and each muscle group. So, there are six models in all for a single athlete. Both of these investigations build gym activity recognition models based on specific exercises and muscle regions. In experiments 1 and 2, each model is programmed to identify all exercises for a particular muscle group. Experiments 3 and 4 are muscle group independent, which contains data from all exercises ignoring the muscle group. Experiment 3 is based on data from all exercises and all the athletes, so we have only one dataset with 44 classes. One model will be trained to identify every activity, regardless of the athlete or muscle area. We have a dataset for every exercise in experiment 4 based on each athlete. One model will be trained to recognize each activity performed by an athlete. In experiments 3 and 4, both models are trained to recognize every exercise on the entire dataset. The Google Colab [35] platform is utilized to conduct the experiments. All four experiments are executed using classification models discussed in the next section.

B. GRID SEARCH CROSS-VALIDATION

There is no method to predict how well a classification model will perform without knowing the value of the hyper-parameters. In order to determine the optimal values, we should ideally try every potential value, which would take time by manual approach. The method of automatically tweaking the hyper-parameter to find the best values for a particular model is known as grid search cross-validation (GridSearchCV). GridSearchCV is a widely recognized technique for hyper-parameter tuning that systematically explores a predefined set of hyper-parameters and evaluates their performance using cross-validation. By leveraging this approach, we aimed to identify the optimal combination of hyper-parameters that balances model complexity and generalizability. GridSearchCV method strives to mitigate both under-fitting and over-fitting risks by applying combination of hyper-parameters. The use of cross-validation ensured that performance is assessed on multiple subsets of the data, thus providing a more accurate estimate of its generalization capabilities.

C. CLASSIFICATION MODELS BASED ON GRIDSEARCHCV

For the gym exercise prediction, we trained the models described below on the transformed train data and then used GridSearchCV to assess how well they performed on the test data. Below is a list of classification models.

1) Decision Tree Classifier (DTC).
2) Gradient Boosted Decision Tree classifier (GB-DT).
3) Linear Support Vector Classifier (L-SVC).
4) Kernel Support Vector Classifier (K-SVC).
5) Logistic Regression (LogReg).
6) Random Forest Classifier (RFC).

V. RESULTS

The outcomes of all the experiments described in the preceding part are presented and examined in this section. All of the experiments are evaluated using the accuracy, precision, recall, and F-measure evaluation criteria. The goal is to recognize the gym exercise that is being performed by the athlete, we have a total of 44 exercises divided into six

TABLE 1. Experiment 1: muscle group dependent accuracy of the classifiers.

Sr	Muscle Group	Classification Algorithm Accuracy (%)						Best Accuracy	Best Classifier
		LogReg	L-SVC	K-SVC	DTC	RFC	GB-DT		
1	Arms	79.70	81.15	83.28	70.03	80.92	81.57	83.28	K-SVC
2	Back	86.49	86.07	91.06	82.04	89.31	88.90	91.06	K-SVC
3	Chest	91.99	91.62	92.23	90.32	92.23	92.27	92.27	GB-DT
4	Legs	80.09	82.55	84.26	80.24	85.98	87.97	87.97	GB-DT
5	Shoulders	69.47	69.59	75.85	64.17	72.70	74.92	75.85	K-SVC
6	CoreBody	91.59	93.35	93.56	88.95	95.18	94.90	95.18	RFC

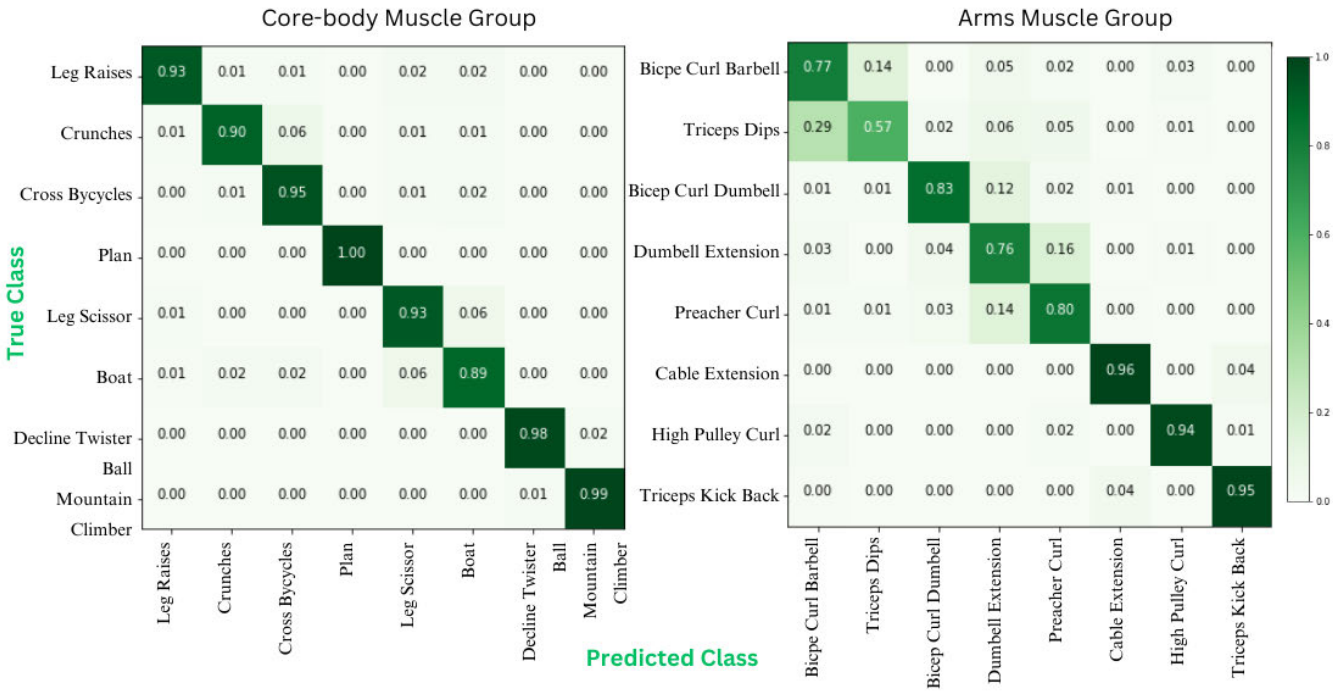

FIGURE 6. Experiment 1: Normalized confusion matrix of core-body and arms muscle group.

muscle groups. Four types of experiments are set up based on muscle group and athlete. These four experiments are tested in two groups. There are two muscle group-dependent experiments and two are muscle group independent. Each of these two groups is further evaluated on the basis of athlete Independence and dependence to analyze the impact of the person performing the exercise as each person may have slight variation while performing an exercise. The following subsections detail the findings from each experiment.

A. EXPERIMENT 1: MUSCLE GROUP DEPENDENT, ATHLETE INDEPENDENT

A different model is trained and tested on the transformed data in this experiment for each muscle group. Each exercise dataset includes information from all of the athletes as well as the class variable's exercise of a single muscle group. In total, we have six datasets for this experiment, one for each muscle group. The accuracy for each classifier and muscle group is depicted in Table 1.

The classification performance based on transformed data for gym exercise recognition on all the muscle groups is good. The random forest classifier performed well on corebody workouts with accuracy above 95.18%. The overall results of gradient boosted decision tree and Kernel SVC with RBF kernel are outstandingly high on all the muscle groups as compared to other algorithms. We achieved an accuracy of above 91% on coreboy, chest, and back, 87.97% on legs, and 83.28% on arms muscle groups. The results of the shoulders muscle group are comparatively low. While few exercises related to shoulders involve less movement of the chest and more movement of arms, still we are able to recognize these exercises with 75.84%. The overall classification performance on only three attributes based on the transformed dataset is outstandingly good on all the muscle groups.

Figure 6 depicts two normalized confusion matrices for the core-body and arms muscle group. The predictions for the best classifier core-body are presented in Figure 6 on left. In the core-body muscle group, the model outperformed with minor miss-classification of leg-scissors for boat exercise and vice versa. Despite the similarity in body posture between both exercises, the arm movements differ slightly. The Plank exercise achieved the best accuracy as it does not involve any movement and athletes have to maintain the same position,

IEEE Access

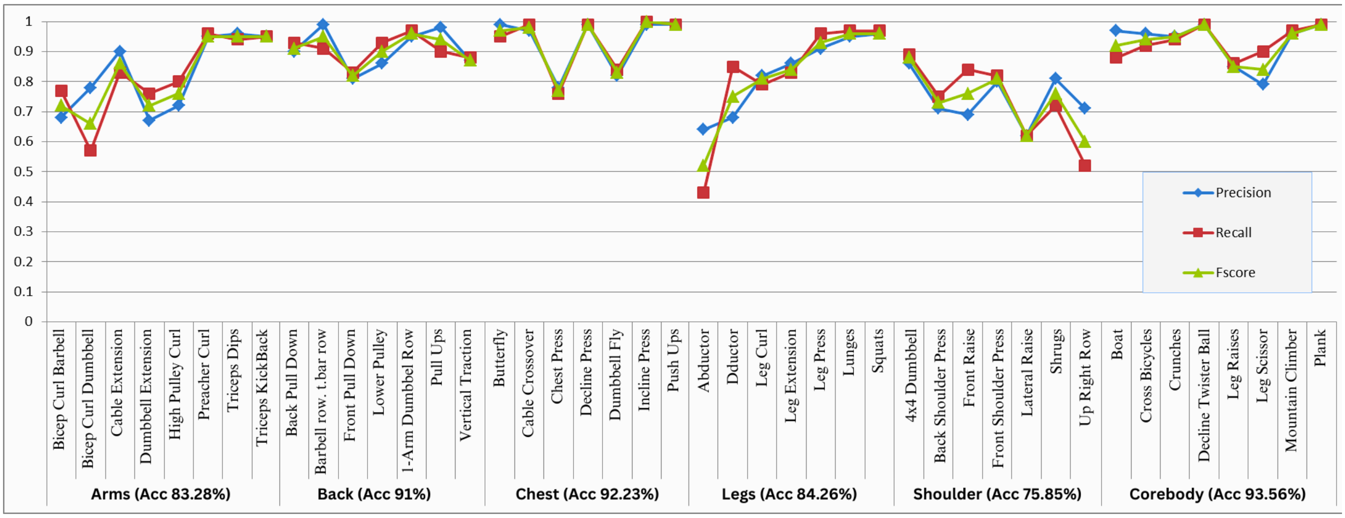

FIGURE 7. Experiment 1: Relevancy for each workout group with the best classifier K-SVC.

TABLE 2. Experiment 2: muscle group and athlete dependent model accuracy of classifiers for each athlete.

Sr	Muscle Group	First Athlete's Classification Accuracy (%)							Second Athlete's Classification Accuracy (%)						
		LogReg	L-SVC	K-SVC	DTC	RFC	GB-DT	Best Classifier	LogReg	L-SVC	K-SVC	DTC	RFC	GB-DT	Best Classifier
1	Arms	80.89	82.51	84.27	70.36	82.11	80.96	K-SVC	84.10	83.93	85.59	74.76	82.88	83.14	K-SVC
2	Back	92.46	92.39	94.55	87.31	92.54	91.57	K-SVC	83.57	84.60	87.23	75.68	84.51	83.85	K-SVC
3	Chest	90.34	90.92	90.59	88.51	90.34	88.51	L-SVC	95.46	95.99	94.93	94.19	95.35	96.09	GD-DT
4	Legs	78.40	79.38	80.42	78.82	85.56	84.79	RFC	87.07	89.12	86.89	80.64	91.26	91.70	RFC
5	Shoulders	67.80	68.53	73.07	60.18	68.29	70.24	K-SVC	78.63	77.61	82.57	70.75	78.56	77.53	K-SVC
6	CoreBody	90.66	91.97	92.05	90.50	93.44	93.67	GD-DT	92.24	93.34	94.89	88.62	94.44	93.60	K-SVC

additionally, there are the highest numbers of instances of this activity as depicted in Figure 1.

Figure 6 presents the normalized confusion matrix for the arms muscle group, the accuracy of which is low as compared to other muscle groups. The exercises cable extension, triceps kick back and high pulley curl have the highest recognition rate with minor misclassification. The triceps dips exercise is the lowest recognition exercise with an accuracy of 0.57%, the reason is that this exercise is done by Sitting on the edge of a weight bench and gripping the edge next to the hips. The movement involved in this exercise is repeatedly going slightly up and down with weight on the arms. This movement does not have much impact on the body posture as posture remains the same in all repetitions.

The precision, recall, and F-score for the best-performing classifier K-SVC according to each muscle group of experiment 1 are depicted in Figure 7. In order to identify gym activity, the transformation of just three accelerometer data attributes for the GAR task is very important. The overall performance of the muscle group-dependent and athlete-independent models is satisfactory.

B. EXPERIMENT 2: ATHLETE AND MUSCLE GROUP DEPENDENT

The experiments in this section are based on the individual muscle group data collected from each participant. For every athlete and muscle area, a different model is developed and put to the test. We are presenting findings for two athletes to validate the performance of the athlete-dependent model. So, in this experiment, we have six datasets organized by muscle group for each athlete, and twelve datasets for two athletes. The classification results on the transformed data for both athletes are depicted in Table 2. The K-SVC algorithm performed well on an athlete-dependent model and achieved the best accuracy for most of the muscle groups and both athletes at the same time. The legs workout is better recognized by Random forest classifiers on both athletes.

TABLE 3. Accuracy comparison of experiment 1 and experiment 2.

Sr	Muscle Group	Classification Accuracy (%)			
		Experiment 2			Experiment 1
		Athlete 1	Athlete 2	Average	All Athletes
1	Arms	84.27	85.59	84.93	83.28
2	Back	94.55	87.23	90.89	91.06
3	Chest	90.92	96.09	93.51	92.27
4	Legs	85.56	91.26	88.41	87.97
5	Shoulders	73.07	82.57	77.82	75.84
6	CoreBody	93.67	94.89	94.28	95.18

A comparison of muscle group-dependent models based on experiments 1 and 2, is presented in Table 3. It can be concluded from the table that there are minor variations of classification accuracy for each athlete in each muscle group, while muscle group-wise results are similar for each athlete. if we look at the average accuracy for

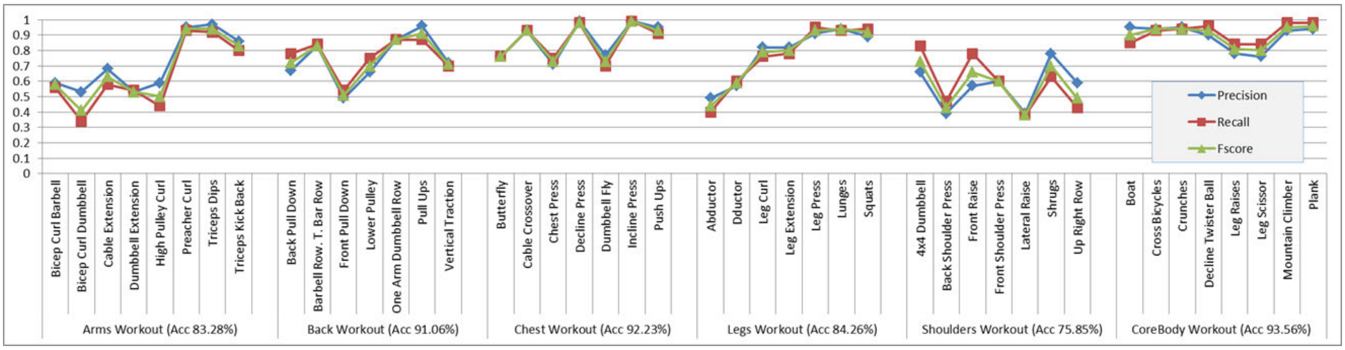

FIGURE 8. Experiment 3: Relevancy for each workout group with the best classifier (K-SVC).

TABLE 4. Experiment 3: classification accuracy for each classifier.

Sr	Classifier	Accuracy
1	Logistic Regression (LogReg)	71.54%
2	Linear Support Vector Classifier (L-SVC)	70.68%
3	Kernel Support Vector Classifier (K-SVC)	76.59%
4	Decision Tree Classifier (DTC)	50.83%
5	Random Forest Classifier (RFC)	73.30%
6	Gradient Boosted Decision Tree classifier	75.14%

athlete-dependent models i.e experiment 2, it is almost the same as athlete-independent models for each muscle group. As the results for the athlete-dependent models are not dominating the athlete-independent model, we can conclude that athlete-independent models are better and will give more ease in further implementations as it involves less dataset.

C. EXPERIMENT 3: MUSCLE GROUP AND ATHLETE INDEPENDENT

A comprehensive dataset, including all athletes, muscle groups, and activities, forms the basis of Experiment 3. All 44 exercises are recognized by a single model that has been trained. The accuracy for each classifier in the muscle group and athlete-independent model is depicted in Table 4. The best classifier for experiment 3 is again the K-SVC with an accuracy of 76.59%. The results of experiment 3 are low as compared to the first two experiments. This is due to the increased number of classes to predict. In experiment 1 three are almost seven classes to predict and in experiment 3 there are 44 classes. Figure 8 summarizes the results of experiment 3 in the form of precision, recall, and fscore for each activity based on the K-SVC classifier's accuracy.

D. EXPERIMENT 4: MUSCLE GROUP INDEPENDENT AND ATHLETE DEPENDENT

Experiment 4 is based on athlete-wise data for all the muscle groups and all exercises. A single model is trained to recognize all the 44 exercises of the athlete. As the results for two athletes are presented in experiment 2, again two athletes results are presented in Table 5 for experiment 4. The best classifier is still the K-SVC with an accuracy of 76.71% for athlete 1 and 81.35% for athlete 2. As we have seen in the experiment 2, that the classification accuracy is higher as compared to athlete 1, similar results are found in this experiment. In comparison with the experiment 3, the results of athlete dependent model are better as compared to athlete independent model as we are dealing with higher number of classes. The muscle group independent results are better when evaluated on each athlete separately.

TABLE 5. Experiment 4: classification accuracy for two athletes.

Sr	Classifier	Accuracy (%)			
		Athlete 1	Athlete 2	Average	Experiment 3
1	LogReg	73.73	77.21	75.47	71.54
2	L-SVC	74.18	77.33	75.76	70.68
3	K-SVC	76.71	81.35	79.03	76.59
4	DTC	58.24	58.71	58.48	50.83
5	RFC	74.09	79.26	76.68	73.30
6	GB-DT	74.24	77.78	76.01	75.14

TABLE 6. Muscle Group dependent accuracy comparison.

Sr	Muscle Group	Accuracy on Feature Engineered Data	Accuracy on Time Series data [32]
1	Arms	83.28%	78%
2	Back	91.06%	88%
3	Chest	92.27%	91%
4	Legs	87.97%	82%
5	Shoulders	75.85%	74%
6	CoreBody	95.18%	90%

E. COMPARATIVE ANALYSIS

This subsection presents a comparative analysis between the proposed feature engineering-based approach and a time series-based approach on the experiment 1. The comparison is conducted with reference to the research presented in [32] on the same dataset in time series format.

The Table 6 presents the accuracy comparison for the muscle group dependent, athlete independent model on both time series and feature engineered data. The accuracy for the arms muscle group improved from 78% on time series data to 83% on feature engineered data. The highest improvement in accuracy is around 6% recorded on the legs muscle group. The accuracy of feature engineering based approach improved among all the muscle groups.

VI. SUMMARY AND CONCLUSION

Gym exercise recognition systems can be used to monitor and manage physical activities in the smart healthcare domain. In this research, we transform raw time-series data to extract new features and identify gym exercises. The collected raw data comprising just three features, we engineered new features from raw data and manage to extract a total of 106 distinctive features. We trained six GridSearchCV based classifiers on the transformed data and evaluate the performance in four types of experimental setups. The featured engineered data improved the overall accuracy over time series data. The classifier K-SVC performs better in most of the scenarios. Among the four types of experiments, the results of experiments 1 and 2 are high with slight variations. Our findings suggest that experiment 1, which is an athlete-independent and muscle group-dependent model, is better. This approach uses data from all athletes, resulting in a consolidated dataset and reduced processing requirements. The results for the muscle group independent experiments 3 and 4 are poor as it involves all the exercises. The feature transformation techniques that we apply in this research are not limited to the gym exercise recognition, but can be extended to any domain involving time-series data including but not limited to sports training and performance analysis, rehabilitation and physical therapy and fitness and health monitoring.

REFERENCES

[1] Y.-L. Hsu, S.-C. Yang, H.-C. Chang, and H.-C. Lai, "Human daily and sport activity recognition using a wearable inertial sensor network," *IEEE Access*, vol. 6, pp. 31715–31728, 2018.

[2] V. Camomilla, E. Bergamini, S. Fantozzi, and G. Vannozzi, "Trends supporting the in-field use of wearable inertial sensors for sport performance evaluation: A systematic review," *Sensors*, vol. 18, no. 3, p. 873, Mar. 2018.

[3] J. C. Quiroz, A. Banerjee, S. M. Dascalu, and S. L. Lau, "Feature selection for activity recognition from smartphone accelerometer data," *Intell. Automat. Soft Comput.*, pp. 1–9, Jul. 2017.

[4] A. Zheng and A. Casari, *Feature Engineering for Machine Learning: Principles and Techniques for Data Scientists*. Sebastopol, CA, USA: O'Reilly Media, 2018.

[5] F. Nargesian, H. Samulowitz, U. Khurana, E. B. Khalil, and D. S. Turaga, "Learning feature engineering for classification," in *Proc. IJCAI*, 2017, pp. 2529–2535.

[6] R. Ahmed Bhuiyan, N. Ahmed, M. Amiruzzaman, and M. R. Islam, "A robust feature extraction model for human activity characterization using 3-axis accelerometer and gyroscope data," *Sensors*, vol. 20, no. 23, p. 6990, Dec. 2020.

[7] A. U. Haq, J. P. Li, S. Ahmad, S. Khan, M. A. Alshara, and R. M. Alotaibi, "Diagnostic approach for accurate diagnosis of COVID-19 employing deep learning and transfer learning techniques through chest X-ray images clinical data in e-healthcare," *Sensors*, vol. 21, no. 24, p. 8219, Dec. 2021.

[8] *Nabriya | Feature Engineering on Time-Series Data*. Accessed: Feb. 10, 2022. [Online]. Available: https://towardsdatascience.com/feature-engineering-on-time-series-data-transforming-signal-data-of-a-smartphone-accelerometer-for-72cbe34b8a6

[9] P. de Souza, D. Silva, I. de Andrade, J. Dias, J. P. Lima, V. Teichrieb, J. P. Quintino, F. Q. B. da Silva, and A. L. M. Santos, "A study on the influence of sensors in frequency and time domains on context recognition," *Sensors*, vol. 23, no. 12, p. 5756, Jun. 2023. [Online]. Available: https://www.mdpi.com/1424-8220/23/12/5756

[10] D. Morris, T. S. Saponas, A. Guillory, and I. Kelner, "RecoFit: Using a wearable sensor to find, recognize, and count repetitive exercises," in *Proc. SIGCHI Conf. Hum. Factors Comput. Syst.*, 2014, pp. 3225–3234.

[11] H. Ding, L. Shangguan, Z. Yang, J. Han, Z. Zhou, P. Yang, W. Xi, and J. Zhao, "FEMO: A platform for free-weight exercise monitoring with RFIDs," in *Proc. 13th ACM Conf. Embedded Netw. Sensor Syst.*, 2015, pp. 141–154.

[12] C. Crema, A. Depari, A. Flammini, E. Sisinni, T. Haslwanter, and S. Salzmann, "IMU-based solution for automatic detection and classification of exercises in the fitness scenario," in *Proc. IEEE Sensors Appl. Symp. (SAS)*, Mar. 2017, pp. 1–6.

[13] C. Shen, B.-J. Ho, and M. Srivastava, "MiLift: Efficient smartwatch-based workout tracking using automatic segmentation," *IEEE Trans. Mobile Comput.*, vol. 17, no. 7, pp. 1609–1622, Jul. 2018.

[14] I. Pernek, K. A. Hummel, and P. Kokol, "Exercise repetition detection for resistance training based on smartphones," *Pers. Ubiquitous Comput.*, vol. 17, pp. 771–782, Dec. 2013.

[15] E. Velloso, A. Bulling, H. Gellersen, W. Ugulino, and H. Fuks, "Qualitative activity recognition of weight lifting exercises," in *Proc. 4th Augmented Hum. Int. Conf.*, 2013, pp. 116–123.

[16] I. Pernek, G. Kurillo, G. Stiglic, and R. Bajcsy, "Recognizing the intensity of strength training exercises with wearable sensors," *J. Biomed. Informat.*, vol. 58, pp. 145–155, Dec. 2015.

[17] P. Hausberger, A. Fernbach, and W. Kastner, "IMU-based smart fitness devices for weight training," in *Proc. 42nd Annu. Conf. IEEE Ind. Electron. Soc. (IECON)*, Oct. 2016, pp. 5182–5189.

[18] E. E. Cust, A. J. Sweeting, K. Ball, and S. Robertson, "Machine and deep learning for sport-specific movement recognition: A systematic review of model development and performance," *J. Sports Sci.*, vol. 37, no. 5, pp. 568–600, Mar. 2019.

[19] H. Ding, J. Han, L. Shangguan, W. Xi, Z. Jiang, Z. Yang, Z. Zhou, P. Yang, and J. Zhao, "A platform for free-weight exercise monitoring with passive tags," *IEEE Trans. Mobile Comput.*, vol. 16, no. 12, pp. 3279–3293, Dec. 2017.

[20] L. Chen, J. Hoey, C. D. Nugent, D. J. Cook, and Z. Yu, "Sensor-based activity recognition," *IEEE Trans. Syst., Man, Cybern., C (Appl. Rev.)*, vol. 42, no. 6, pp. 790–808, Nov. 2012.

[21] J. R. Kwapisz, G. M. Weiss, and S. A. Moore, "Activity recognition using cell phone accelerometers," *ACM SIGKDD Explor. Newslett.*, vol. 12, no. 2, pp. 74–82, Mar. 2011.

[22] J. Margarito, R. Helaoui, A. M. Bianchi, F. Sartor, and A. G. Bonomi, "User-independent recognition of sports activities from a single Wrist-Worn accelerometer: A template-matching-based approach," *IEEE Trans. Biomed. Eng.*, vol. 63, no. 4, pp. 788–796, Apr. 2016.

[23] E. Mitchell, D. Monaghan, and N. O'Connor, "Classification of sporting activities using smartphone accelerometers," *Sensors*, vol. 13, no. 4, pp. 5317–5337, Apr. 2013.

[24] A. Godfrey, A. K. Bourke, G. M. Ólaighin, P. van de Ven, and J. Nelson, "Activity classification using a single chest mounted tri-axial accelerometer," *Med. Eng. Phys.*, vol. 33, no. 9, pp. 1127–1135, Nov. 2011.

[25] M. Nazarahari and H. Rouhani, "Detection of daily postures and walking modalities using a single chest-mounted tri-axial accelerometer," *Med. Eng. Phys.*, vol. 57, pp. 75–81, Jul. 2018.

[26] M. Ermes, J. Pärkkä, J. Mäntyjärvi, and I. Korhonen, "Detection of daily activities and sports with wearable sensors in controlled and uncontrolled conditions," *IEEE Trans. Inf. Technol. Biomed.*, vol. 12, no. 1, pp. 20–26, Jan. 2008.

[27] B. Fu, L. Jarms, F. Kirchbuchner, and A. Kuijper, "ExerTrack—Towards smart surfaces to track exercises," *Technologies*, vol. 8, no. 1, p. 17, Mar. 2020.

[28] M. Cornacchia, K. Ozcan, Y. Zheng, and S. Velipasalar, "A survey on activity detection and classification using wearable sensors," *IEEE Sensors J.*, vol. 17, no. 2, pp. 386–403, Jan. 2017.

[29] Y. Wang, S. Cang, and H. Yu, "A survey on wearable sensor modality centred human activity recognition in health care," *Expert Syst. Appl.*, vol. 137, pp. 167–190, Dec. 2019.

[30] Z. Hussain, M. Sheng, and W. Emma Zhang, "Different approaches for human activity recognition: A survey," 2019, *arXiv:1906.05074*.

[31] A. Hussain, K. Zafar, and A. R. Baig, "Fog-centric IoT based framework for healthcare monitoring, management and early warning system," *IEEE Access*, vol. 9, pp. 74168–74179, 2021.

[32] A. Hussain, K. Zafar, A. R. Baig, R. Almakki, L. AlSuwaidan, and S. Khan, "Sensor-based gym physical exercise recognition: Data acquisition and experiments," *Sensors*, vol. 22, no. 7, p. 2489, Mar. 2022. [Online]. Available: https://www.mdpi.com/1424-8220/22/7/2489

[33] *Medtronic ZephyrPerformance Systems | Performance Monitoring Technology*. Accessed: Feb. 10, 2022. [Online]. Available: https://www.zephyranywhere.com/
[34] *Government College University Faisalabad*. Accessed: Jun. 6, 2023. [Online]. Available: https://www.gcuf.edu.pk/
[35] G. Colab. *Welcome To Colaboratory*. Accessed: Feb. 10, 2022. [Online]. Available: https://colab.research.google.com/

AFZAAL HUSSAIN received the Ph.D. degree from the National University of Computers and Emerging Sciences, Islamabad, Pakistan. With an illustrious academic journey, he currently holds the position of an Assistant Professor with the Information Technology Department, Government College University Faisalabad, Pakistan, since 2012. His scholarly pursuits are rooted in cutting-edge domains, with a particular focus on machine learning, big data, and data mining. His expertise and dedication exemplify his commitment to advancing knowledge and innovation in the ever-evolving field of information technology.

MUHAMMAD ADEEL ZAHID received the Ph.D. degree from the Department of Computer Science, Lahore University of Management Sciences, Pakistan. He is currently a Lecturer in computer science with Government College University Faisalabad, Pakistan. His research interests include distributed systems, service-oriented computing, and machine learning.

USAMA AHMED received the M.S. degree in information security from the National University of Science and Technology (NUST), Islamabad, and the Ph.D. degree from COMSATS University Islamabad, Lahore Campus, specializing in trust-aware cloud systems. He was the Project Manager in the security and cryptology cell (SCC) with Air University, Islamabad, for information security planning and architecture research for next generation networks. He is currently an Assistant Professor with the Department of Software Engineering, Government College University Faisalabad. He has numerous publications on his account in conferences and impact factor journals. His research interests include the IoT and cloud security, risk management, and trust-aware distributed systems.

SHAHBAZ NAZEER received the B.S. (Hons.) and M.S. degrees in computer science from Government College University Faisalabad, Pakistan, in 2007 and 2011, respectively. Currently, he is pursuing the Ph.D. degree in computer science with Agriculture University, Faisalabad, Pakistan. He has been a Lecturer with the Information Technology Department, Government College University Faisalabad, since 2008. He is currently an accomplished academic professional with a distinguished academic background. His research interests include machine learning and data mining, reflecting his commitment to advancing knowledge, and innovation in computer science. During the B.S. degree, he was awarded the Gold Medal for his exceptional performance.

KASHIF ZAFAR has been associated with the Department of Computer Science, National University of Computer and Emerging Sciences, since 2004. Prior to joining NUCES, he was a Faculty Member with the Faculty of Computer Science and Engineering, GIKI, from 2002 to 2004, specializes in the field of artificial intelligence. He has worked in the areas of machine learning, swarm intelligence, computational intelligence, evolutionary algorithms, data mining, and expert systems. He has been involved in teaching and research at undergraduate and graduate levels and has supervised numerous final-year projects and M.S. thesis. He has been awarded the Full Scholarship from the Foundation for the Advancement of Science and Technology, Islamabad, for his postgraduate studies. He has published a number of research papers in international journals and conferences of high repute. He has hands-on experience in the IT industry and has worked in different organizations in USA and has around 15 years of full-time job experience.

ABDUL RAUF BAIG (Member, IEEE) received the B.E. degree in electrical engineering from the NED University of Engineering and Technology, Karachi, Pakistan, in 1987, the Diplómes de Spécialisation degree in computer science from Supélec, Rennes, France, in 1996, and the Ph.D. degree in computer science from the University of Rennes 1, Rennes, in 2000. He was with the National University of Computer and Emerging Sciences, Islamabad, Pakistan, as a Faculty Member (an Assistant Professor, an Associate Professor, and a Professor) from 2001 to 2010, and is currently on leave. Since 2010, he has been a Professor with Imam Mohammad Ibn Saud Islamic University, Riyadh, Saudi Arabia. He has more than 100 publications in journals and international conferences. His current research interests include machine learning, data mining, and evolutionary algorithms. He is a member of the IEEE Computational Intelligence Society.