IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 14 August 2023, accepted 25 September 2023, date of publication 28 September 2023, date of current version 6 October 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3320144

RESEARCH ARTICLE

LogRF: An Approach to Human Pose Estimation Using Skeleton Landmarks for Physiotherapy Fitness Exercise Correction

ALI RAZA[1], AZAM MEHMOOD QADRI[1], IQRA AKHTAR[2], NAGWAN ABDEL SAMEE[3], AND MAALI ALABDULHAFITH[3]

[1]Institute of Computer Science, Khwaja Fareed University of Engineering and Information Technology, Rahim Yar Khan 64200, Pakistan
[2]Department of Electrical and Biomedical Engineering, Khwaja Fareed University of Engineering and Information Technology, Rahim Yar Khan 64200, Pakistan
[3]Department of Information Technology, College of Computer and Information Sciences, Princess Nourah bint Abdulrahman University, Riyadh 11671, Saudi Arabia

Corresponding author: Nagwan Abdel Samee (nmabdelsamee@pnu.edu.sa)

This work was supported by Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia, through Researchers Supporting under Project PNURSP2023R407.

ABSTRACT Human pose and gesture estimation are crucial in correcting physiotherapy fitness exercises. In recent years, advancements in computer vision and machine learning approaches have led to the development of sophisticated pose estimation models that accurately track and analyze human movements in real time. This technology enables physiotherapists and fitness trainers to gain valuable insights into their client's exercise forms and techniques, facilitating more effective exercise corrections and personalized training regimens. This research aims to propose an efficient artificial intelligence method for human pose estimation during physiotherapy fitness exercises. We utilized a multi-class exercise dataset based on human skeleton movement points to conduct our experimental research. The dataset comprises 133 features derived from human skeleton movements during various exercises, resulting in high feature dimensionality that affects the performance of human pose estimation with machine learning and deep learning methods. We have introduced a novel Logistic regression Recursive Feature elimination (LogRF) method for feature selection. Extensive experiments demonstrate that using the top twenty selected features, the random forest method outperformed state-of-the-art studies with a high-performance score of 0.998. The performance of each applied method is validated through a k-fold approach and further enhanced using hyperparameter tuning. Our proposed study assists specialists in identifying and addressing potential biomechanical issues, improper postures, and incorrect movement patterns, which are essential for injury prevention and optimizing exercise outcomes. Furthermore, this study enhances the capabilities of remote monitoring and guidance capabilities, allowing physiotherapists to support their patient's progress with prescribed exercises continually.

INDEX TERMS Machine learning, deep learning, human pose, gesture estimation, physiotherapy, skeleton landmarks, feature engineering.

I. INTRODUCTION

Physiotherapy plays a crucial role in enhancing an individual's overall well-being and rehabilitation through targeted fitness exercises [1]. In recent years, integrating machine learning techniques has shown great promise in assisting physiotherapists with accurate, real-time feedback during exercise sessions [2]. Human pose and gesture estimation have emerged as critical components in this domain, enabling the automatic and precise analysis of patient's movements. By employing sophisticated algorithms, researchers have endeavoured to detect and track key body joints and gestures, facilitating the identification of incorrect postures and movements during exercises [3]. However, neglecting physiotherapy fitness exercises has various adverse consequences, including an increased risk of developing chronic diseases

The associate editor coordinating the review of this manuscript and approving it for publication was Tao Liu.

and higher mortality rates. Regular physical activity, particularly tailored to individual needs through physiotherapy, has been proven to be an effective preventive measure against chronic conditions such as cardiovascular diseases, diabetes, and obesity [4]. Failure to engage in adequate physical exercise can lead to a sedentary lifestyle, promoting weight gain and metabolic dysregulation, significantly contributing to the onset and progression of these debilitating diseases.

Human pose and gesture estimation using Artificial Intelligence (AI) methods has emerged as a cutting-edge technology with immense potential in physiotherapy fitness exercise correction based on skeleton landmarks [5], [6]. AI systems can accurately and dynamically analyze patient's movements during exercise sessions by harnessing the power of computer vision and machine learning techniques. Combined with deep learning models, these sophisticated algorithms enable detecting and tracking key body joints and gestures, providing real-time feedback to physiotherapists and patients [7]. Using AI for Human Pose and Gesture Estimation offers numerous advantages in physiotherapy, such as automatic and precise posture and movement analysis, leading to improved exercise form and reduced risk of injury [8]. With the capability to identify incorrect postures and movements, AI-based solutions empower physiotherapists to provide personalized corrective guidance, optimizing the rehabilitation process and promoting faster recovery. Moreover, these innovative technologies facilitate remote monitoring, allowing patients to access physiotherapy sessions from the comfort of their homes and encouraging compliance with exercise regimens.

This research introduces an AI-based mechanism for human pose estimation during physiotherapy fitness exercises. A multi-class exercise dataset comprising 133 features derived from human skeleton movements during various exercises is utilized to develop applied machine learning and deep learning methods. The dataset's high feature dimensionality impacts the performance of human pose estimation. We propose a novel LogRF method for feature selection. The LogRF method aims to eliminate the least important features progressively. This is accomplished through iterative training of the logistic regression model, wherein the feature with the smallest weight magnitude is eliminated in each iteration.

Our main research contributions toward human pose and gesture estimation are as follows:

- A new LogRF method is proposed for selecting features in human pose estimation during physiotherapy fitness exercises. The proposed LogRF method selects the top twenty features through an iterative mechanism. In each iteration, the feature with the smallest weight magnitude is eliminated.
- For performance comparison, we employed four advanced machine learning and deep learning approaches. The k-fold cross-validation approach is utilized to validate the performance, and hyperparameter tuning is implemented to optimize it. Furthermore, a computational complexity analysis is also conducted.

The remaining research study is followed as Section II, which elaborates on the literature analysis. Section III describes our study methodology, while Section IV evaluates the results of the applied methods. Our main findings are described in Section V.

II. LITERATURE ANALYSIS

This literature analysis section aims to explore various studies, academic papers, and techniques employed in human pose and gesture estimation using machine learning approaches. The analysis delves into the evolution of this technology, highlighting key advancements and breakthroughs over the years. The section also discusses the strengths and limitations of these approaches, comparing their performance metrics, accuracy, and scalability.

In this study [9], the authors present a system that recognizes multiple poses of yoga exercises performed by trainees. The new data is created using an HD 720P web camera from different locations through video recording for testing, which contains six different yoga pose records from videos performed by fifteen individuals. The system consists of two main phases: data extraction from the media pipe and preprocessing of the data for training and testing using a classification based on five machine-learning approaches. The feature engineering process is conducted during preprocessing. The system achieved an accuracy score of 94% using the logistic regression model, outperforming the other classified machine learning models.

In this study [10], the author assesses several human poses through computer vision using real-time and recorded videos. Using computer vision, RGB-quality images were utilized to detect human postures, with different human body parts and graphics serving as input data for assessing human posture. Thirty-three pose milestones were observed using OpenCV and Mediapipe in this research. The BlazePose GHUM 3D Pose Landmark Model was employed to evaluate the 2D human pose, achieving an accuracy of 96.9% with visibility points.

In this study [11], the author presents the emerging and exciting scope of AI in assessing the pose of humans. This research was specially developed for bowler's pose estimation. The dataset used in this study was novel, which the author collected by recording different bowlers' actions and movements via an HD video camera. A deep learning technique was proposed to estimate the bowlers' pose, classify, and assess players. The suggested deep learning model for bowling (BowlingDL) and MoveNet models were utilized to estimate the poses of different bowlers. The proposed model attained an 80% accuracy score on training data and an 83% accuracy score on testing data. An innovative mobile device application, constructed for the bowlers with BowlingDL standard performance using lite TensorFlow, was deployed.

In this paper [12], the author asserts that yoga is an effective and disciplined physical activity for enhancing body muscle strength and overall fitness. This proposed study utilizes an interactive system that simultaneously detects six yoga poses

TABLE 1. The employed literature summary analysis.

Ref.	Year	Dataset	Learning Type	Proposed Technique	Accuracy
[9]	2017	Self-made dataset by HD cam	Machine learning	Logistic Regression	94.00%
[12]	2018	New Data recorded by Kinet sensor.	Machine learning	Adaboost	94.78%
[13]	2019	SYSU 3D HOI and HUA datasets.	Deep learning	MDTW	93.13%
[14]	2020	ETRI dataset	Deep learning	InceptionResNetV2	95.34%
[15]	2020	Yoga-82S, 82 instance	Deep learning	DenseNet	91.44%
[16]	2022	New Data collected by Kinect sensor	Deep learning	CNN-LSTM	90.18%
[17]	2022	Coco and MPII human Pose dataset	Deep learning	PoseNet	97.60%
[10]	2023	RGB image dataset	Deep learning	BlazePose GHUM 3D	96.90%
[11]	2023	Self-collected dataset by HD cam	Deep learning	MoveNet	83.00%

performed by six individuals, providing real-time guidance through voice cues, visual guides, and pose images. To create the database for pose detection, the system employed the Adaboost algorithm and utilized a software development kit specifically designed for the Kinect sensor. An expert yoga trainer conducted the data collection process, and the accuracy achieved in detecting yoga poses was 94.78%.

The author of this study [16] presents that in the past, human activity recognition has been an exciting topic utilized in many previous studies in healthcare and human intervention [18], [19], [20]. Multiple AI models were employed for activity recognition, but they all failed to select optimal features for long-term human activity recognition. A new dataset from 20 participants using the Kinect V2 sensor containing 12 classes of human physical activities was developed to address this. A hybrid technique combining Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM) was introduced for activity identification. CNN is used to extract attributes, while LSTM is used to learn information. To achieve the best solution for human activity recognition, a comprehensive study was conducted using standard machine learning and deep learning algorithms. The results were promising, with the CNN-LSTM approach obtaining a 90.18% accuracy score, demonstrating its suitability for human activity recognition.

In this paper [17], the author asserts that pose estimation in humans is highly effective in multiple fields, such as producing movies involving multiple human actions and enabling object movements in video games. Human poses can also be utilized to execute user interfaces on mobile devices. This study employs four pose estimation models for contrast and classification: OpenPose, PoseNet, MoveNet Lightning, and MoveNet. The Coco and MPII Human Pose datasets are used in this study to compare the performance of these models for pose estimation. The models' performances are evaluated, resulting in the following accuracy scores: OpenPose achieves 86.2% accuracy, PoseNet achieves 97.6% accuracy, MoveNet achieves 75.1% accuracy, and MoveNet Thunder achieves 80.6% accuracy. PoseNet, our proposed model, demonstrates superior performance with a 97.6% accuracy score for estimating poses on mobile devices.

This paper's author [14] contributes to posture identification using ensemble CNN in the home atmosphere. This study is specially designed for senior citizens who live alone at home. Posture identification is essential for helping senior people to avoid risks. The study's analysis uses a pose dataset developed by the Korean Electronics and Telecommunications Research Institute (ETRI). The dataset is collected from 51 home circumstances with ten poses, containing 51,000 recorded images. The authors employed a deep learning approach to achieve the desired outcomes from the image data. Five preprocessing approaches were employed for investigations: VGGNet, ResNet, DenseNet, InceptionResNet, and Xception, based on trained CNN. The analysis demonstrates that InceptionResNetV2s provide excellent performance with an average accuracy score of 95.34%, outperforming the other trained CNNs.

In computer vision, human pose assessment is a widely identified challenge that involves determining the positions of joints [15]. However, the previous dataset used in the research was not authentic enough to address the issues of pose diversity, object occlusion, and viewpoints adequately. In this study, the author proposes the dataset yoga-82S for large-scale pose identification with 82 classes, introducing the idea of pose assessment as a classification task. The dataset is collected from various websites and comprises a hierarchical structure consisting of body positions, variations in body positions, and corresponding pose names. The classification accuracy of advanced CNN architectures on Yoga-82 is demonstrated. The researchers also introduce several hierarchical adaptations of DenseNet tailored to the hierarchical labels in the dataset, effectively improving performance compared to other classification methods. Specifically, in the 3rd level class, DenseNet-201 achieved an accuracy of 74.91%, while DensNet-121 obtained an impressive 91.44% accuracy score.

The authors in [13] present a framework designed for families using fog computing in this paper. It relies on three phases: joint mobility assessment, investigating the abnormality of actions of upper limbs, and abnormal gait detection for lower limbs. They introduce a semi-automatic approach for evaluating upper limb motion called Rapid Upper Limb Assessment (RULA) using Kinect v2. The study uses the standard 3D action dataset and the Human Upper Action dataset (HUA) for experiments. The Semi-automatic Rapid Upper Limb Assessment (RULA) using Kinect v2 approach is employed to evaluate upper limb motion. The modified Dynamic Time Warping (DTW) algorithm was evaluated on the HUA dataset, resulting in an accuracy of 89.50%. The RULA dataset is utilized for gait abnormality detection, achieving an excellent accuracy of 93.13% with the Modified DTW model.

III. PROPOSED METHODOLOGY

This section examines our proposed methodology for human pose estimation during various exercises. The methods employed for analysis and result calculations are comprehensively discussed. We present a step-by-step exploration of our proposed methodology in this section.

Figure 1 illustrates the architectural workflow analysis of our proposed research methodology. We utilized a publicly available dataset for experiments involving human poses during various exercises. Initially, the dataset exhibited high-dimensional features for experimentation. To address this, we introduced a novel approach for feature selection, focusing on retaining only those features that significantly contribute to human pose estimation. The selected feature set is then divided into training (80%) and testing (20%) subsets. We proceeded to train and test several advanced AI approaches. The performance of the hyperparameter-tuned models is evaluated using unseen test data. The method that demonstrated superior performance is subsequently employed for human pose estimation, specifically for correcting exercises in physiotherapy fitness routines.

A. MULTI-CLASS EXERCISE POSES FOR HUMAN SKELETON

This study utilized a multi-class exercise dataset [21] based on human skeleton pose data to conduct our experimental research. The dataset comprises 2701 rows and 133 columns, with features derived from human skeleton movements during various exercises. Each row corresponds to a specific exercise, while each column represents different aspects of the human skeleton model. The dataset includes coordinates for the X, Y, and Z axes and visibility values for 33 landmarks, resulting in 132 values per exercise. To analyze various exercise poses, the dataset encompasses seven distinct target classes: "Rest" (406 rows), "Left Bicep" (435 rows), "Right Bicep" (369 rows), "Left Shoulder" (373 rows), "Right Shoulder" (401 rows), "Left Tricep" (317 rows), and "Right Tricep" (399 rows), as illustrated in Figure 2.

B. NOVEL FEATURE ELIMINATION

Our novel proposed feature elimination approach, LogRF, is analyzed in this section. We input the original dataset with 133 features into our proposed approach. The working architectural analysis of the proposed feature selection is illustrated in Figure 3. The LogRF approach identifies a subset of features that contribute the most to the model's predictive power, aiding in dealing with a large number of features. When combined with the logistic regression method, the proposed LogRF iteratively selects and eliminates features to find the most informative subset for optimal model performance. The selected features with high importance lead to improved model generalization and performance.

Furthermore, we conducted a feature space analysis on twenty carefully chosen features. The results of this analysis are depicted in Figure 4. This analysis clearly illustrates the high separability of the selected features, ultimately contributing to the strong performance in human pose estimation.

1) PROPOSED LOGRF MATHEMATICAL ALGORITHM

Given a skeleton landmarks data with N samples and M features ($M = 133$ in our case), the logistic regression model can be represented as:

$$\hat{y}_i = \frac{1}{1 + e^{-\mathbf{w}^T \mathbf{x}_i}} \tag{1}$$

where $\hat{y}_i$ is the predicted probability for the i-th sample, $\mathbf{w}$ is the weight vector, and $\mathbf{x}_i$ is the feature vector.

The objective of the LogRF method is to eliminate the least important features recursively. This is achieved by iteratively training the logistic regression model and eliminating the feature with the smallest weight magnitude. The process can be summarized as follows:

1) Train the logistic regression method on the current set of features.
2) Calculate the absolute magnitude of the weight vector: $|\mathbf{w}|$.
3) Identify the feature with the smallest weight magnitude: $j = \arg\min |\mathbf{w}|$.
4) Remove the feature with index j from the dataset.
5) Repeat the process until the desired number of features is reached.

C. APPLIED ARTIFICIAL INTELLIGENCE TECHNIQUES

In recent years, integrating applied AI techniques for human pose and gesture estimation in physiotherapy has shown significant potential for improving fitness exercise correction [22], [23], [24]. This approach enables real-time and accurate analysis of human movement during exercises by employing advanced deep-learning algorithms and computer vision methods [25]. The AI-powered system can provide valuable feedback to physiotherapists and patients, facilitating precise form adjustments, reducing the risk of injuries, and optimizing exercise efficacy.

1) RANDOM FOREST

The Random Forest (RF) method performs classification tasks using multiple decision trees [26]. In the context of human pose and gesture estimation based on skeleton landmarks, we can represent the RF algorithm as follows:

The basic equation for the decision tree model is given by:

$$\hat{y}_i = f(x_i) = \sum_{m=1}^{M} c_m \cdot I(x_i \in R_m) \tag{2}$$

where:

- $\hat{y}_i$ is the forecast output for the i-th data sample,
- x_i represents the input features (skeleton landmarks) for the i-th data sample,

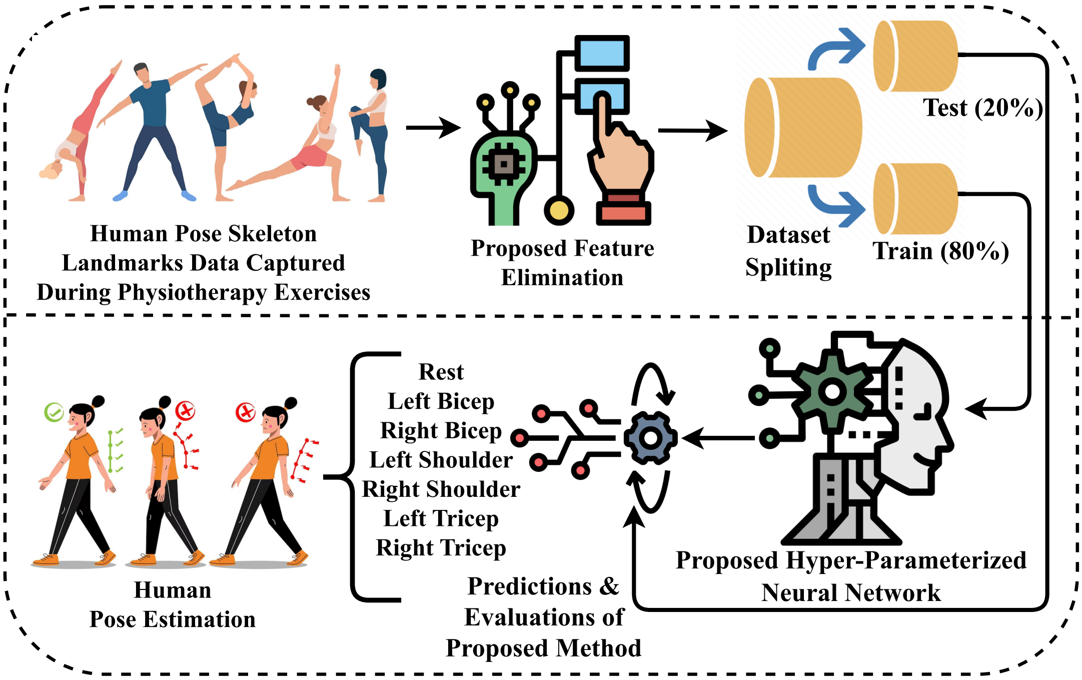


FIGURE 1. Architectural analysis of our proposed research methodology.

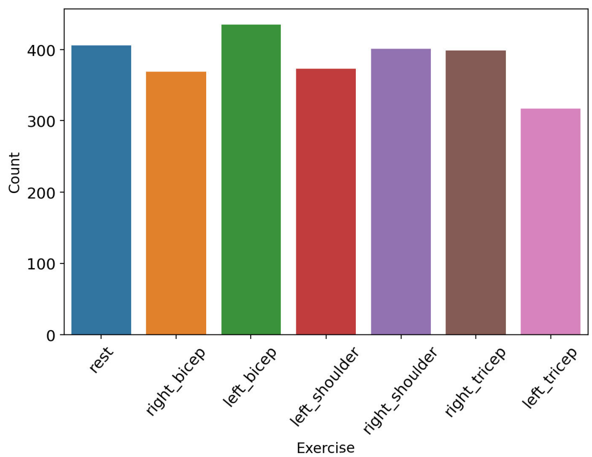


FIGURE 2. The target distributions analysis for multi-class exercise poses data.

- $f(x_i)$ is the decision tree's prediction function for the input x_i,
- M is the number of leaf nodes (terminal nodes) in the decision tree,
- c_m is the output value attached with the m-th leaf node, and
- $I(x_i \in R_m)$ is an indicator function that returns 1 if the input x_i falls into the m-th leaf node (region) R_m, and 0 otherwise.

2. For the RF method, we combine multiple decision trees to form an ensemble, and the final prediction is obtained by averaging the predictions of all the individual trees:

$$\hat{y}_{\mathrm{RF}}(x_i) = \frac{1}{N_{\mathrm{trees}}} \sum_{j=1}^{N_{\mathrm{trees}}} f_j(x_i) \tag{3}$$

where:

- $\hat{y}_{\mathrm{RF}}(x_i)$ is the RF's prediction for the input x_i,
- N_{trees} is the total number of decision trees in the Random Forest, and
- $f_j(x_i)$ is the prediction of the j-th decision tree for the input x_i.

2) LOGISTIC REGRESSION

Logistic Regression (LR) is a classification approach commonly used in machine learning for tasks like human pose and gesture estimation [27]. The method LR is used to transform the linear union of input features and corresponding weights into the probability:

$$P(y = 1|\mathbf{X}) = \frac{1}{1 + e^{-\mathbf{X}^T\mathbf{W}}} \tag{4}$$

where:

- $\mathbf{X}$ is the input data feature vector of size $1 \times N$, where N is the number of total features.
- $\mathbf{W}$ is the weight vector of size $N \times 1$ containing the parameters of the LR model.
- $\mathbf{X}^T$ denotes the transpose of $\mathbf{X}$.
- e is Euler's number and contains approximately 2.71828 value.

3) GATED RECURRENT UNIT

The Gated Recurrent Unit (GRU) [28] is a popular recurrent neural network architecture used for sequential data processing, such as human pose and gesture estimation based on skeleton landmarks.

The GRU consists of three main gates: reset gate (r_t), update gate (z_t), and the candidate hidden state (h_t). The

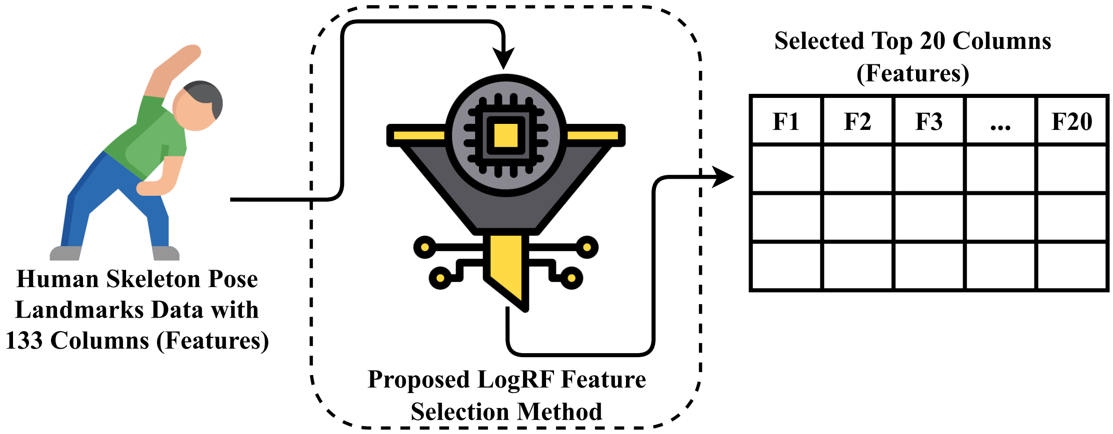


FIGURE 3. The proposed feature selection architecture analysis.

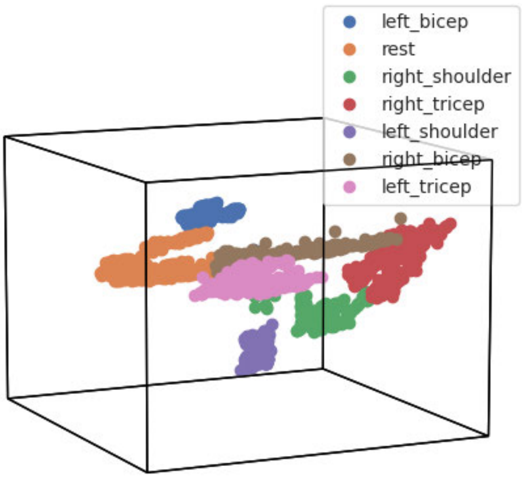


FIGURE 4. Feature space analysis of selected dataset features using the proposed approach.

equations for updating these gates at each time step t are as follows:

The reset gate is as follows:

$$r_t = \sigma(W_r \cdot [h_{t-1}, x_t]) \tag{5}$$

The update gate is as follows:

$$z_t = \sigma(W_z \cdot [h_{t-1}, x_t]) \tag{6}$$

Candidate's hidden state is as follows:

$$\tilde{h}_t = \tanh(W \cdot [r_t \odot h_{t-1}, x_t]) \tag{7}$$

where:

h_{t-1} is the previous hidden state,
x_t represents the input at value time step t,
W_r, W_z, W are learnable weight matrices,
σ is the sigmoid activation function,
$[a, b]$ denotes the combine of vectors a and b,
$\odot$ represents element-wise multiplication.

Finally, the updated hidden state h_t is computed as a linear interpolation in between the earlier hidden state h_{t-1} and the candidate hidden state $\tilde{h}_t$ using the update gate z_t:

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t \tag{8}$$

This recurrent updating process is applied iteratively over the entire sequence of skeleton landmarks, allowing the GRU to capture temporal dependencies and extract meaningful pose and gesture information.

4) LONG SHORT-TERM MEMORY

The Long Short-Term Memory (LSTM) network [29] is a kind of recurrent neural network that can effectively model sequential data such as human pose and gesture information based on skeleton landmarks. In an LSTM unit, the following equations govern its operation:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \tag{9}$$
$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{10}$$
$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \tag{11}$$
$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t \tag{12}$$
$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \tag{13}$$
$$h_t = o_t \odot \tanh(C_t) \tag{14}$$

where:

i_t, f_t, o_t are the input, forget, and output gate at a time step t,
$\tilde{C}_t$ is the candidate cell state at time step t,
C_t is the cell state value at time step t,
h_t is the hidden state value at time step t,
x_t is the input value at time step t,
$[h_{t-1}, x_t]$ denotes the combination of h_{t-1} and x_t,
σ is the sigmoid activation function, and
$\odot$ represents element-wise multiplication.

The LSTM network's ability to capture temporal dependencies and handle long-term memory makes it suitable for human pose and gesture estimation tasks based on sequential skeleton landmarks data.

TABLE 2. Hyperparameter setting analysis of applied method in this study.

Method	Hyperparameter Description
RF	n_estimators=4, max_depth=2, random_state=0, criterion='entropy'
LR	random_state=0, max_iter=12, multi_class='auto', C=1.0
GRU	Dropout=0.2, activation='softmax', loss = 'categorical_crossentropy', optimizer = 'adam' , metrics='accuracy'
LSTM	Dropout=0.2, activation='softmax', loss = 'categorical_crossentropy', optimizer = 'adam' , metrics='accuracy'

D. HYPERPARAMETER OPTIMIZATION

The most appropriate parameters for the applied deep learning and machine learning algorithms are selected using a recursive testing and training process [30]. Table 2 demonstrates the best-fit selected parameters for our applied human pose and gesture assessment models.

IV. RESULTS AND DISCUSSIONS

This section provides a thorough analysis of the outcomes obtained from employing machine learning techniques in the field of human pose estimation. Within this section, we examine the empirical findings derived from our experiments and thoroughly explore their implications. The results presented highlight the precision and effectiveness of each algorithm, as measured by various performance metrics.

A. EXPERIMENTAL SETTING

To conduct our research experiments on human pose estimation, we have developed advanced AI approaches using the Python 3.0 programming language. We employed an open-source environment called Google Colab to implement our experiments. This environment features a GPU backend, 13 GB of RAM, and 90 GB of disk space. We evaluated the performance of the applied pose estimation method using metrics such as precision, accuracy, recall, f1 score, standard deviations, and runtime computation.

B. RESULTS WITH ORIGINAL FEATURES

This section comprises the performance analysis using all features. Each applied machine learning and deep learning model undergoes evaluation by utilizing all dataset features, as illustrated in Table 3. The parameters employed to assess the results of the applied methods include accuracy, recall, and F1 score. Additionally, the classification report is analyzed. The analysis reveals that machine learning models outperform deep learning methods. Among all dataset features, the LSTM method yields the lowest performance score, followed by the GRU approach. The machine learning-based RF achieves an acceptable performance score of 87%. Conversely, the LR approach attains the highest accuracy score of 95% in comparison to the others. This analysis concludes that the method exhibits low performance for human pose and gesture estimation when all dataset features are applied. Consequently, there exists a requirement for an advanced feature selection mechanism to mitigate dimensionality and enhance performance.

TABLE 3. Performance metrics analysis results with original features.

Method	Accuracy	Target class	Precision	Recall	F1
RF	0.87	left_bicep	1.00	1.00	1.00
		left_shoulder	1.00	1.00	1.00
		left_tricep	0.00	0.00	0.00
		rest	1.00	1.00	1.00
		right_bicep	0.53	1.00	0.69
		right_shoulder	1.00	1.00	1.00
		right_tricep	1.00	1.00	1.00
		Average	0.80	0.87	0.83
LR	0.95	left_bicep	1.00	0.90	0..95
		left_shoulder	0.89	0.98	0.93
		left_tricep	1.00	1.00	1.00
		rest	1.00	1.00	1.00
		right_bicep	0.80	1.00	0.89
		right_shoulder	1.00	0.77	0.87
		right_tricep	1.00	1.00	1.00
		Average	0.96	0.95	0.95
GRU	0.64	left_bicep	0.52	0.99	0.68
		left_shoulder	0.53	0.85	0.65
		left_tricep	0.64	0.28	0.39
		rest	0.88	0.49	0.63
		right_bicep	0.96	0.84	0.90
		right_shoulder	0.60	0.70	0.65
		right_tricep	0.90	0.40	0.55
		Average	0.71	0.64	0.63
LSTM	0.22	left_bicep	0.10	0.35	0.15
		left_shoulder	0.00	0.00	0.00
		left_tricep	0.28	0.31	0.29
		rest	0.00	0.00	0.00
		right_bicep	0.00	0.00	0.00
		right_shoulder	0.38	0.82	0.52
		right_tricep	0.00	0.00	0.00
		Average	0.11	0.22	0.14

C. RESULTS WITH NOVEL PROPOSED APPROACH

Using the proposed LogRF approach, the twenty selective features are utilized to assess the performance of the applied methods as described in Table 4. The deep learning-based LSTM method achieved an accuracy score of 0.829, which is acceptable and significantly improved compared to the results with the original features. The GRU approach also performs well, with a performance accuracy of 0.974. The machine learning models demonstrate strong performance with the selected features. The highest performance accuracy of 0.998 is achieved by the proposed RF approach. Additionally, the classification report results for each method show significant improvement. This analysis concludes that all applied methods enhanced their performance accuracy with these highly important selective features. Our proposed approach proves to be beneficial in enhancing human pose estimation performance.

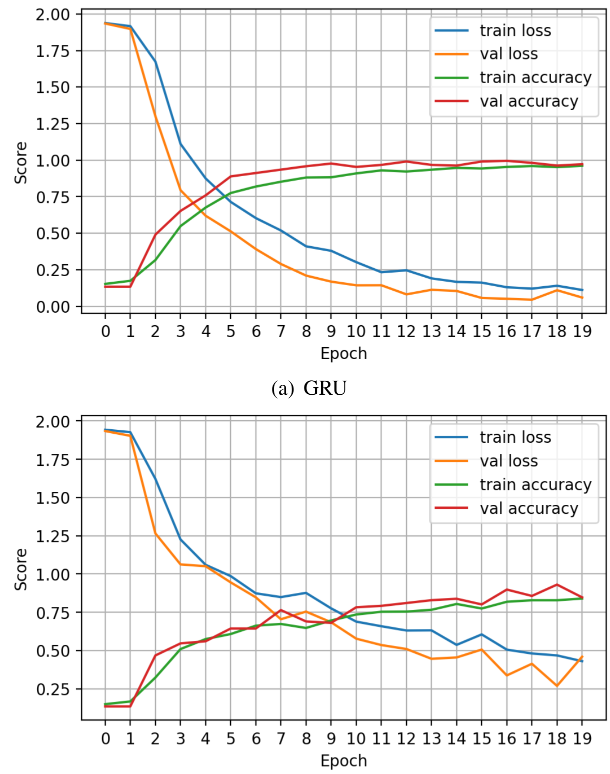
The performance analysis of applied neural network techniques based on time series is presented in Figure 5. The training loss, validation loss, training accuracy, and validation accuracy are the performance metrics evaluated during the training process of GRU and LSTM models. The analysis indicates that in the initial five epochs, loss scores are high, and accuracy scores are low due to the random weight initialization of the neural network at the start layer. After achieving optimal weights in the sixth epoch, both deep learning methods improve their performance scores and reduce the loss. This analysis concludes that the deep learning models

TABLE 4. Performance metrics analysis results with selective features.

Method	Accuracy	Target class	Precision	Recall	F1
RF	**0.998**	left_bicep	1.00	1.00	1.00
		left_shoulder	1.00	0.98	0.99
		left_tricep	1.00	1.00	1.00
		rest	0.99	1.00	0.99
		right_bicep	1.00	1.00	1.00
		right_shoulder	1.00	1.00	1.00
		right_tricep	1.00	1.00	1.00
		Average	**1.00**	**1.00**	**1.00**
LR	0.990	left_bicep	1.00	1.00	1.00
		left_shoulder	1.00	0.98	0.99
		left_tricep	0.99	1.00	0.99
		rest	1.00	1.00	1.00
		right_bicep	0.99	0.96	0.97
		right_shoulder	0.96	0.99	0.98
		right_tricep	1.00	1.00	1.00
		Average	0.99	0.99	0.99
GRU	0.974	left_bicep	0.96	1.00	0.98
		left_shoulder	0.98	1.00	0.99
		left_tricep	0.99	0.90	0.94
		rest	1.00	1.00	1.00
		right_bicep	1.00	0.97	0.98
		right_shoulder	0.90	0.95	0.93
		right_tricep	1.00	1.00	1.00
		Average	0.97	0.97	0.97
LSTM	0.829	left_bicep	0.62	1.00	0.76
		left_shoulder	0.99	0.99	0.99
		left_tricep	0.89	0.37	0.53
		rest	1.00	0.90	0.94
		right_bicep	0.83	0.92	0.87
		right_shoulder	0.80	0.84	0.81
		right_tricep	0.85	0.83	0.84
		Average	0.86	0.83	0.82

attained commendable scores of 90 and above when trained on selected dataset features.

The performance validation analysis, based on the confusion matrix, of the applied methods with selective features is depicted in Figure 6. The analysis reveals that machine learning techniques exhibited a lower error rate during classification compared to deep learning approaches. Specifically, the LSTM model produced 65 incorrect predictions, indicating a significant error rate. In contrast, the proposed RF approach yielded only 1 incorrect prediction in the context of human pose estimation, thereby confirming the exceptional performance of the proposed model.



(a) GRU

(b) LSTM

FIGURE 5. Time series-based performance comparison analysis of applied neural network approaches during training.

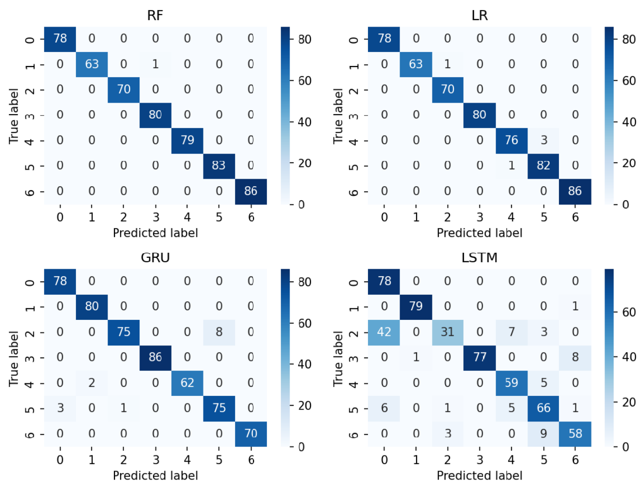


FIGURE 6. Confusion matrix analysis of the applied approaches.

D. K-FOLD CROSS VALIDATIONS ANALYSIS

We utilized k-fold cross-validation mechanisms to validate the performance of each applied method, as outlined in Table 5. The cross-validation analysis is based on k-fold accuracy and standard deviations. The dataset of selective features is divided into ten folds and used for evaluation. The analysis reveals that deep learning models exhibited low-performance scores with high standard deviations during validations. In comparison, the machine learning method achieved favourable validation scores. This analysis concludes that the proposed RF method attained a high k-fold accuracy score of 0.99 with a minimal standard deviation of 0.0042. Our proposed method has been successfully validated in a generalized manner for human pose estimations during fitness exercise corrections.

TABLE 5. Performance validation analysis results with proposed features.

Method	k-fold	Accuracy	Standard deviations (+/-)
RF	**10**	**0.99**	**0.0042**
LR	10	0.98	0.0054
GRU	10	0.84	0.1259
LSTM	10	0.68	0.2149

E. COMPUTATIONS COMPLEXITY ANALYSIS

The computational complexity analysis of the applied methods is presented in Table 6. We calculated the runtime computation scores in seconds for each applied method during dataset construction. Upon comparison, it was found

TABLE 6. Computations complexity analysis with proposed features.

Method	Runtime Computations (seconds)
RF	**0.033**
LR	0.062
GRU	10.47
LSTM	12.62

TABLE 7. Performance comparisons with pose estimation based state of the art studies.

Ref	Proposed Technique	Performance Accuracy
[31]	Ensemble Model	0.93
[32]	Scalable Neural Network	0.95
[33]	Two Stream Bilinear C3D	0.84
Our Study	**LogRF+RF**	**0.99**

that deep learning models exhibit the highest complexity, which subsequently results in lower performance. The machine learning-based LR achieved lower complexity when compared to the deep learning methods. The analysis demonstrates that the proposed RF method achieved the minimum computation score of 0.033 seconds for human pose estimation.

F. STATE OF THE ART COMPARISON

The performance of the proposed method was compared with state-of-the-art pose estimation studies, as outlined in Table 7. We selected advanced pose estimation studies for this comparison. The analysis demonstrates that the state-of-the-art studies achieved moderate performance scores. In contrast, our proposed approach outperformed these studies, achieving a high accuracy score of 0.99 for human pose estimation.

V. CONCLUSION AND FUTURE WORK

This research proposes an efficient artificial intelligence method for human pose estimation during physiotherapy fitness exercises. We utilized a multi-class exercise dataset based on human skeleton movement points to conduct our experimental research. The dataset has high feature dimensionality, which affects the performance of human pose estimation with machine learning and deep learning methods. In this research, we have introduced a novel LogRF method for feature selection. We employed four advanced machine learning and deep learning approaches for performance comparison. Extensive experiments demonstrate that the RF method outperformed state-of-the-art studies using the top twenty selected features with a high-performance score of 0.998. The results of each applied method are validated through a k-fold approach and further enhanced using hyperparameter tuning.

A. FUTURE WORK

For future directions, we plan to create an interactive graphical user interface. This interface will incorporate our proposed backend system and connect to a real-time camera. The intended interface aims to accurately estimate human poses for the purpose of correcting physiotherapy fitness exercises.

ACKNOWLEDGMENT

This work was funded by Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2023R407), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia. The authors would like to express their gratitude to Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2023R407), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

REFERENCES

[1] K. Wittmeier and K. Mulder, "Enhancing lifestyle for individuals with haemophilia through physical activity and exercise: The role of physiotherapy," *Haemophilia*, vol. 13, no. s2, pp. 31–37, Sep. 2007.
[2] D. Deotale, M. Verma, P. Suresh, and N. Kumar, "Physiotherapy-based human activity recognition using deep learning," *Neural Comput. Appl.*, vol. 35, no. 15, pp. 11431–11444, May 2023.
[3] O. Stamm and L. Perotti, "Expert requirements for an ultrasound-based wearable using deep learning for exercise feedback in older chronic back pain patients," in *Proc. Int. Conf. Hum.-Comput. Interact.* Cham, Switzerland: Springer, 2022, pp. 158–173.
[4] M. Yokogawa, Y. Taniguchi, and Y. Yoneda, "Qualitative research concerning physiotherapy approaches to encourage physical activity in older adults with dementia," *PLoS ONE*, vol. 18, no. 7, Jul. 2023, Art. no. e0289290.
[5] K.-C. Lin and R.-J. Wai, "A feasible fall evaluation system via artificial intelligence gesture detection of gait and balance for sub-healthy community-dwelling older adults in Taiwan," *IEEE Access*, vol. 9, pp. 146404–146413, 2021.
[6] R. Abdulghafor, A. Abdelmohsen, S. Turaev, M. A. Ali, and S. Wani, "An analysis of body language of patients using artificial intelligence," *Healthcare*, vol. 10, no. 12, p. 2504, 2022.
[7] G. Samhitha, D. S. Rao, C. Rupa, Y. Ekshitha, and R. Jaswanthi, "Vyayam: Artificial intelligence based bicep curl workout tacking system," in *Proc. Int. Conf. Innov. Comput., Intell. Commun. Smart Electr. Syst. (ICSES)*, Sep. 2021, pp. 1–5.
[8] J. J. Ojeda-Castelo, M. D. L. M. Capobianco-Uriarte, J. A. Piedra-Fernandez, and R. Ayala, "A survey on intelligent gesture recognition techniques," *IEEE Access*, vol. 10, pp. 87135–87156, 2022.
[9] U. Bahukhandi and S. Gupta, "Yoga pose detection and classification using machine learning techniques," *Int. Res. J. Mod. Eng. Technol. Sci.*, vol. 3, no. 12, pp. 13–15, 2021.
[10] M. Shah, K. Gandhi, B. M. Pandhi, P. Padhiyar, and S. Degadwala, "Computer vision & deep learning based realtime and pre-recorded human pose estimation," in *Proc. 2nd Int. Conf. Appl. Artif. Intell. Comput. (ICAAIC)*, May 2023, pp. 313–319.
[11] N. F. Janbi and N. Almuaythir, "BowlingDL: A deep learning-based bowling players pose estimation and classification," in *Proc. 1st Int. Conf. Adv. Innov. Smart Cities (ICAISC)*, Jan. 2023, pp. 1–6.
[12] E. W. Trejo and P. Yuan, "Recognition of yoga poses through an interactive system with Kinect based on confidence value," in *Proc. 3rd Int. Conf. Adv. Robot. Mechatronics (ICARM)*, Jul. 2018, pp. 606–611.
[13] W. Cao, J. Zhong, G. Cao, and Z. He, "Physiological function assessment based on Kinect v2," *IEEE Access*, vol. 7, pp. 105638–105651, 2019.
[14] Y.-H. Byeon, J.-Y. Lee, D.-H. Kim, and K.-C. Kwak, "Posture recognition using ensemble deep models under various home environments," *Appl. Sci.*, vol. 10, no. 4, p. 1287, Feb. 2020.
[15] M. Verma, S. Kumawat, Y. Nakashima, and S. Raman, "Yoga-82: A new dataset for fine-grained classification of human poses," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jun. 2020, pp. 4472–4479.
[16] I. U. Khan, S. Afzal, and J. W. Lee, "Human activity recognition via hybrid deep learning based model," *Sensors*, vol. 22, no. 1, p. 323, Jan. 2022.
[17] B. Jo and S. Kim, "Comparative analysis of OpenPose, PoseNet, and MoveNet models for pose estimation in mobile devices," *Traitement du Signal*, vol. 39, no. 1, pp. 119–124, Feb. 2022.
[18] A. Raza, H. U. R. Siddiqui, K. Munir, M. Almutairi, F. Rustam, and I. Ashraf, "Ensemble learning-based feature engineering to analyze maternal health during pregnancy and health risk prediction," *PLoS ONE*, vol. 17, no. 11, Nov. 2022, Art. no. e0276525.

[19] A. Raza, F. Rustam, H. U. R. Siddiqui, I. D. L. T. Diez, and I. Ashraf, "Predicting microbe organisms using data of living micro forms of life and hybrid microbes classifier," *PLoS ONE*, vol. 18, no. 4, Apr. 2023, Art. no. e0284522.

[20] A. Raza, F. Rustam, H. U. R. Siddiqui, I. D. L. T. Diez, B. Garcia-Zapirain, E. Lee, and I. Ashraf, "Predicting genetic disorder and types of disorder using chain classifier approach," *Genes*, vol. 14, no. 1, p. 71, Dec. 2022.

[21] D. Patel. (May 2023). *Multi-Class Exercise Poses for Human Skeleton | Kaggle*. Accessed: Sep. 8, 2023. [Online]. Available: https://www.kaggle.com/datasets/dp5995/gym-exercise-mediapipe-33-landmarks

[22] L. D. Haberkamp, M. C. Garcia, and D. M. Bazett-Jones, "Validity of an artificial intelligence, human pose estimation model for measuring single-leg squat kinematics," *J. Biomech.*, vol. 144, Nov. 2022, Art. no. 111333.

[23] A. D. Pazho, "Ancilia: Scalable intelligent video surveillance for the artificial intelligence of things," *IEEE Internet Things J.*, vol. 10, no. 17, pp. 14940–14951, Sep. 2023.

[24] A. Raza, F. Rustam, B. Mallampati, P. Gali, and I. Ashraf, "Preventing crimes through gunshots recognition using novel feature engineering and meta-learning approach," *IEEE Access*, early access, Sep. 18, 2023, doi: 10.1109/ACCESS.2023.3316695.

[25] A. Raza, K. Munir, M. S. Almutairi, and R. Sehar, "Novel class probability features for optimizing network attack detection with machine learning," *IEEE Access*, vol. 11, pp. 98685–98694, 2023.

[26] Z. Feng, Y. Shi, D. Zhou, and L. Mo, "Research on human activity recognition based on random forest classifier," in *Proc. IEEE Int. Conf. Control, Electron. Comput. Technol. (ICCECT)*, Apr. 2023, pp. 1507–1513.

[27] S. Alqadhi, J. Mallick, S. Talukdar, A. A. Bindajam, T. K. Saha, M. Ahmed, and R. A. Khan, "Combining logistic regression-based hybrid optimized machine learning algorithms with sensitivity analysis to achieve robust landslide susceptibility mapping," *Geocarto Int.*, vol. 37, no. 25, pp. 9518–9543, Dec. 2022.

[28] A. Raza, K. Munir, M. Almutairi, F. Younas, M. M. S. Fareed, and G. Ahmed, "A novel approach to classify telescopic sensors data using bidirectional-gated recurrent neural networks," *Appl. Sci.*, vol. 12, no. 20, p. 10268, Oct. 2022, Art. no. 10268.

[29] A. R. Inturi, V. M. Manikandan, and V. Garrapally, "A novel vision-based fall detection scheme using keypoints of human skeleton with long short-term memory network," *Arabian J. Sci. Eng.*, vol. 48, no. 2, pp. 1143–1155, Feb. 2023.

[30] A. Raza, M. R. Al Nasar, E. S. Hanandeh, R. A. Zitar, A. Y. Nasereddin, and L. Abualigah, "A novel methodology for human kinematics motion detection based on smartphones sensor data using artificial intelligence," *Technologies*, vol. 11, no. 2, p. 55, Apr. 2023.

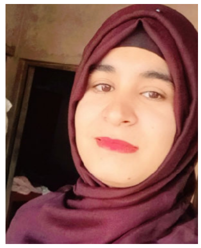
[31] H. Luo, M. Wang, P. K.-Y. Wong, and J. C. P. Cheng, "Full body pose estimation of construction equipment using computer vision and deep learning techniques," *Autom. Construct.*, vol. 110, Feb. 2020, Art. no. 103016.

[32] J. Liang and M. Lin, "Shape-aware human pose and shape reconstruction using multi-view images," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2019, pp. 4351–4361.

[33] C. Cao, Y. Zhang, C. Zhang, and H. Lu, "Body joint guided 3-D deep convolutional descriptors for action recognition," *IEEE Trans. Cybern.*, vol. 48, no. 3, pp. 1095–1108, Mar. 2018.

ALI RAZA received the Bachelor of Science degree in computer science from the Department of Computer Science, Khwaja Fareed University of Engineering and Information Technology (KFUEIT), Rahim Yar Khan, Pakistan, in 2021, and the M.S. degree in computer science from KFUEIT. He has published several articles in reputed journals. His current research interests include data science, artificial intelligence, data mining, natural language processing, machine learning, deep learning, and image processing.

AZAM MEHMOOD QADRI received the master's degree in information technology from the Department of Computer Science and Information Technology, Virtual University, Lahore, Pakistan, in 2017. He is currently pursuing the M.S. degree in computer science with the Khwaja Fareed University of Engineering and Information Technology (KFUEIT), Rahim Yar Khan, Pakistan. His research interests include machine learning, deep learning, and artificial intelligence.



IQRA AKHTAR received the Bachelor of Science degree in physics from the Institute of Physics, Khwaja Fareed University of Engineering and Information Technology (KFUEIT), Rahim Yar Khan, Pakistan, where she is currently pursuing the M.S. degree in electronics with the Department of Electrical and Biomedical Engineering. Her current research interests include artificial intelligence, physics applications, and physical science.

NAGWAN ABDEL SAMEE received the B.S. degree in computer engineering from Ein Shams University, Egypt, in 2000, and the M.S. degree in computer engineering and the Ph.D. degree in systems and biomedical engineering from Cairo University, Egypt, in 2008 and 2012, respectively. Since 2013, she has been an Assistant Professor with the Information Technology Department, CCIS, Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia. Her research interests include data science, machine learning, bioinformatics, and parallel computing. Her awards and honors include the Takafull Prize (Innovation Project Track), Princess Nourah Award in innovation, the Mastery Award in predictive analytics (IBM), the Mastery Award in Big Data (IBM), and the Mastery Award in Cloud Computing (IBM).

MAALI ALABDULHAFITH was born in Saudi Arabia, in September 1985. She received the Doctor of Philosophy (Ph.D.) degree in computer science from Dalhousie University, Halifax, Canada, in 2017. In 2014, she joined the College of Computer and Information Science (CCIS), Princess Nourah bint Abdulrahman University (PNU), as a Lecturer. She was promoted to an Assistant Professor, in 2018. Currently, she is the Director of Data Management and Performance Measurement with CCIS, PNU, overlooking and managing the strategy of the college. Her research interests include machine learning, data analytics, emerging wireless technology, and technology applications in health care.