IJWR

Human Activity Recognition Using a Hybrid Approach of Radial Basis Neural Networks and Support Vector Machines

Fatemeh Azimzadeh [a]*, Ali Ghoroghi [b]
[a] Scientific Information Database, ACECR, Tehran, Iran; azimzadeh@sid.ir
[b] School of Engineering, Cardiff University, Cardiff, UK; ghoroghi@cardiff.ac.uk

ABSTRACT

The Internet of Things (IoT) has become increasingly prevalent, and recent advances in machine learning, particularly in healthcare, have gained significant attention from researchers. One prominent interdisciplinary topic in these fields is human activity recognition (HAR). Despite extensive research, several challenges remain in this area, especially concerning the application of modern machine learning techniques for HAR. This study proposes a novel method for human activity recognition by combining radial basis function neural networks (RBFNN) and support vector machines (SVM). The approach enhances recognition accuracy and algorithm efficiency by extracting relevant features using RBFNN and convolutional neural networks (CNN). Classification is then performed using SVM. The proposed method was evaluated using the UCI HAR dataset, which includes six distinct human activities. Results demonstrate that the proposed approach achieves an accuracy of 99%, surpassing existing methods.

*Keywords***— *Human Activity Recognition, Radial Basis Neural Network Algorithm, Support Vector Machine Algorithm***

1. Introduction

The rapid development of electronics is rendering wearable devices an integral part of everyday human life. A key service these devices offer is the continuous monitoring of users' movements and activities. This service is made possible with the help of accurate internal accelerometers which can easily provide users with information about their movements and activities 24 hours a day, 7 days a week. Although health and fitness are two of the main areas where wearable devices are commonly applied, continuous monitoring of activities and movements with such devices has potential applications in many other areas, including smart living and indoor positioning [1]. On the other hand, semantic challenges have become increasingly significant in the field of the Internet of Things [2]. In spite of the ever-surging demand for wearable devices, battery recharging is regarded as one of the major obstacles to more widespread employment of these devices. Battery technology has advanced considerably in recent years; most wearable devices, however, still need to be recharged daily or at least once a week. One of the existing trends to decrease dependence on battery recharging is known as kinetic energy harvesting which involves converting kinetic energy of human activities and movements to electrical energy that can be consumed by wearable devices. Energy extracted from wearable devices contains patterns which are essential for the recognition of human activities.

To recognize human activities, the present study used kinetic energy data extracted from wearable devices. Noteworthy is the fact that this study did not research issues related to kinetic energy harvesting hardware and only concentrated on the use of extracted data for the recognition of human activities. As such, the problem addressed regards the recognition of human activities using data extracted from kinetic energy. This is a categorization problem and, thus, needs to be solved using an appropriate classification algorithm. One of the main shortcomings of the designs of previous research studies is the low degrees of their classification accuracy.

doi http://dx.doi.org/10.22133/ijwr.2025.491948.1250
Citation F. Azimzadeh, A. Ghoroghi, "Human Activity Recognition Using a Hybrid Approach of Radial Basis Neural Networks and Support Vector Machines", *International Journal of Web Research*, vol.8, no.1,pp.49-59, 2025, doi: http://dx.doi.org/10.22133/ijwr.2025.491948.1250.

*Coressponding Author
Article History: Received: 2 September 2024 ; Revised: 1 December 2024; Accepted: 11 December 2024.

Methods for recognizing human activities can be classified into three categories: methods based on machine vision, environmental methods, and methods based on wearable devices. Recognition of human activities based on machine vision has been the focus of research for a long time and has nearly reached maturity. To mention the main problems with these methods, one can refer to the costly installation and maintenance of the cameras, together with ignoring privacy and ethical issues [3]. With respect to the second class of methods, that is environmental methods for recognizing human activities, it can be held that through such methods outcomes are affected negatively due to the existence of high exposure to noise and interference caused by the movements of people and other objects. Additionally, these methods can be applied to only recognize the activities that entail interactions between humans and objects [4]. Considering the above-mentioned points, experts may argue that activity recognition through wearable devices is the most reliable of these three methods. As stated previously, the most important challenge of these devices is battery life, which is currently deemed to be even a more serious concern given the users' increasing demands for more power and efficiency. Although much work has been done to enhance battery technology to meet user demands, batteries are still one of the biggest restricting factors in developing wearable technology [5]. The principal innovation of this research study is presenting a new method for recognizing human activities through a combination of support vector machine and radial basis neural networks. This method, while involving low computational complexity, tries to obviate the need for sensors and at the same time increase the accuracy of human activity recognition.

2. Previous Studies

All classification techniques seek to solve an optimization problem. As such, they can be divided into two categories: generative and discriminative. Generative techniques consider probabilistic patterns according to specific parameters between data and classes and determine a common distribution or model for the specified data and classes. This can be a direct model or a conditional distribution of data using Bayesian law. Generative techniques aim to estimate basic parameters and use them to update the data classification. On the other hand, in discriminative techniques, it is assumed that there is a certain distance and similarity criterion between the two patterns. In other words, there is a certain similarity between the data samples of every class which are different from the data samples of other classes. In general, it can be maintained that discriminative techniques are memory-free and non-parametric.

Whereas generative techniques are less popular because of the high costs they incur for the patterns, discriminative techniques are more frequently employed, especially in cases where the amount of data is substantial. As one of the main classification techniques, discriminative techniques are used to recognize human activities. These techniques, in turn, include various procedures and methods such as decision trees, hierarchical threshold, fuzzy logic, clustering, artificial neural networks, and support vector machines. Owing to the complexity and high volumes of calculations inherent in some classification techniques, it is not possible to teach classification online and, therefore, it is performed offline, which makes the activity recognition a time-consuming process. Recognizing activities in real time is still a major challenge in this area. Moreover, the issue of differences in the patterns of human activities is also inevitable in everyday life, and classification techniques should pay more attention to this issue. With regard to the points and facts mentioned above, no classifier can be definitively recommended for the recognition of activities since each technique pays special attention to some of the challenges in this area and tries to solve them. Therefore, arriving at an appropriate classifier (with a high accuracy, appropriate timing, coverage of all simple and complex activities, etc.) to identify human activities in the real world is still an important area of research [6].

In [7], machine learning method was used as a logical model for predicting human activities and movements from inertia sensors installed on smartphones. This study was intended to demonstrate the learning capabilities of the logical model machine in obtaining higher prediction rates even when fed with short-time sections of data (one second) compared to the longer time sections (5.2 seconds) used in other studies. The performance of the system designed based on logical model machine learning was compared with the performance of a system designed based on a random forest and a logical regression tree for a set of static and dynamic activities. The system was trained and tested on two publicly available datasets referred to as WISDM and UCI HAR. The proposed logical model machine learning method outperformed RF and LR by achieving 86.90% and 94.02% recognition accuracy on WISDM and UCI HAR, respectively. In addition, this system obtained between 82.89% and 88.73% overall accuracy during the evaluation of the reciprocal dataset.

In [8], a supervised learning method is described. This method was devised to recognize human activities based on data collected from human movements. The principal challenge when working with HAR involves overcoming the problems which arise with the fixed and moving nature of activity signals. This study proposes a categorization model based on a two-channel convolutional neural network which uses the frequency and power features of the collected human action signals. The proposed model was tested on the UCI HAR dataset, which resulted in a categorization accuracy of 95.25%.

Xia [9] suggests a deep neural network which combines convolutional layers with long-term memory. The model presented in this study can automatically extract activity features and categorize them with reference to several model parameters. It should be stated that LSTM is a type of recursive neural network which is more suitable for processing time series. In the proposed architecture of this model, the raw data collected were fed by moving sensors to a convolution layer. Furthermore, a Global Average Pooling (GAP) layer was employed to completely replace a post-convolutional layer so as to reduce model parameters. Moreover, a Batch Normalization (BN) layer was added after the gap layer in an attempt to accelerate convergence, which led to obtaining tangible results. The performance of this model was assessed on three general datasets: UCI HAR, WISDM, and OPPORTUNITY. The assessment results revealed that the overall accuracy of the model was 95.78% in the UCI HAR dataset, 95.85% in the WISDM dataset, and 92.63% in the OPPORTUNITY dataset. The findings also indicated that the proposed model enjoys better resistance and activity recognition capabilities than the reported results. It follows that this model is not only capable of extracting the features of activities, but it also has fewer parameters and higher accuracy.

Dua [10] presents a deep neural network model using a convolutional neural network and the Gated Recurrent Unit (GRU) as an ultimate model for automatic feature extraction and activity classification. This research project conducted tests using raw data which were obtained from wearable sensors with nominal pre-processing. Thus, it did not include any machine feature extraction techniques. More specifically, it should be noted that this model first used CNN to extract features. This technique extracted spatial features which were given in the form of time series properties as input to the GRU. The GRU used two gates, Reset (Reset Gate) and Update (Update Gate), to map the extracted features to hidden vectors, which were in fact the same time series features extracted from the input data. The extracted features were next given to fully-connected layers, and these layers were used for feature learning. In the multi-input mode, the inputs were given to several convolutions simultaneously, and the output of the convulsions ended up in GRUs, which played a feature-extracting role. The accuracy obtained on UCI HAR, PAMAP2, and WISDM datasets was 96.20%, 97.21% and 95.27%, respectively.

Li et al. [11] introduced a HAR algorithm, HAR_WCNN, which utilizes a wide time-domain convolutional neural network and multi-environment sensor data for daily activity recognition. Challa et al. [12] proposed a hybrid model combining a convolutional neural network (CNN) with bidirectional long short-term memory (BiLSTM). This multi-branch CNN-BiLSTM network performs automatic feature extraction from raw sensor data with minimal preprocessing. The integration of CNN and BiLSTM enables the model to learn both local features and long-term dependencies in sequential data. Kaya and Topuz [13] developed a sensor-based activity recognition method using a 1D-CNN deep learning approach to detect human activities.

Aburuyesh et al. [14] proposed a wireless sensing technique that leverages Channel State Information (CSI) for behavior recognition. This method enhances accuracy while using low-cost resources and enables through-wall and wide-angle predictions using WiFi signals. Similarly, Chen et al. [15] introduced Wisor-DL, a lightweight HAR system that reconstructs WiFi CSI tensors and applies deep learning for improved recognition.

Feng et al. [16] designed an unsupervised domain adaptation framework, Adversarial Time-Frequency Attention (ATFA), to efficiently adapt models to new users. The attention-based modality fusion module selectively captures and integrates important modalities based on their context, reducing redundant information.

3. Research Method

Recognizing human activities remains a major challenge. In recent years, several deep learning methods, with different areas of application, have been proposed to enhance diagnostic performance. The present research study aimed to combine the methods of support vector machine and a convolutional and radial basis neural network with two different approaches in order to replace the previous methods. To this aim, dimension and feature reduction was carried out using two procedures: Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA). Subsequently, the outputs of these procedures were combined to extract the features using convolutional and radial basis neural networks. Next, the extracted target features were fed as input to the vector machine algorithm so that the SVM classification algorithm would be performed and the category type would be selected. Furthermore, all the above-mentioned steps were performed without using the methods of dimension and feature reduction, and the outputs were fed to the algorithm. Also, to compare with basic algorithms such as SVM, KNN, and RandomForest, the researchers performed the above steps through both approaches (with dimension reduction and without dimension reduction), the results of which are available. Figure 1 shows the general research method.

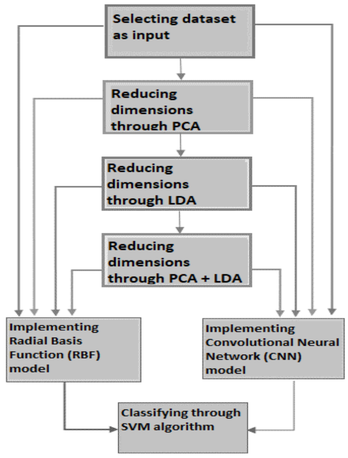

Figure. 1. The General Research Method

In the proposed method, a convolutional neural network and a radial basis neural network were used in conjunction with one another, which allowed the possibility of having stronger features extracted from sequence frames. The resulting feature vector was entered as input to the support vector machine classification with a radial basis neural network kernel. This was done in order to assign each sample to the corresponding tag and recognize the activity performed. In other words, the selected dataset was first fed as input to the radial basis network algorithm where the target features were extracted and, in turn, entered as input to the vector machine algorithm so that the SVM algorithm could classify and select the category type. Subsequently, the pre-processing of the dataset, which consisted of four steps, was applied to the data and then it was considered as input for the proposed model. The four steps of the proposed method were as follows:

(First step) Reduction of model dimensions: Input datasets had considerably large dimensions, and this volume of data tends to decrease the possibility of achieving the desirable high accuracy in a typical model. For this reason, two algorithms (PCA and LDA) were used to reduce the model dimensions.

(Second step) Normalization: The Rescaling (Min-Max Normalization) method was applied.

(Third step) Determination of training and test data: For this purpose, 70% of the dataset was selected for training, and 30% for testing.

(Fourth step) Conversion of labels into one-hot encoding: A one-hot vector was considered for each data tag. This vector had one cell of 1 and n-1 cells of zero.

After the fourth step, the data was transferred to deep learning and branching network models. In their last layer, these networks had fully-connected layers, which decreased the accuracy of the model. Consequently, the researchers used a support vector machine algorithm. The steps of the algorithm were as follows:

1. Execution of branching neural network: This step involved training the branching network on the dataset.
2. Finishing features: All fully-connected layers were cut and removed from the branch network structure. The location cutting point was then used as the feature selection.
3. Implementation of the SVM model: The features selected in the third step were given to the SVM. This was done in an attempt to help learn the data better with the RBF kernel.
4. Finishing labeling: Once the training process of the SVM was completed, the test data was given to it.
5. Identification of activity type: In this step, the prediction results were mapped to one of the data categories which became one-hot vectors in the pre-processing step.

3.1. Data Collection

The UCI HAR dataset consisted of 30 volunteers ranging in age from 19 to 48 years. For each person, six activities (walking, walking upstairs, walking downstairs, sitting, standing, lying) were collected by having them wear a Samsung Galaxy S II smartphone on their back. The accelerometer and the built-in gyroscope recorded the 3-axis linear acceleration and the 3-axis angular velocity at a constant rate of 50 Hz. Tests for the manual labeling of the data were videotaped. The resulting dataset had 7352 records and 561 fields. For each record in this dataset, the following specifications were collected and are available through the link[1]:

– three-axis acceleration of the accelerometer (total acceleration) and estimated body acceleration,
– three-axis angular velocity of the gyroscope,
– a 561-feature vector with time and frequency variables,
– activity tag, and
– the ID of the subject who performed the test.

[1]https://archive.ics.uci.edu/ml/datasets/human+activity+recognition+using+smartphones

3.2. Reduction of the Dataset Dimensions

In general, the larger the dimensions (number of features) of the problem being explored, the more scattered the records in the search space. To solve this problem, the researchers selected a subset of features. Through this procedure, the features whose data value was low for exploration were selected to be removed. For this reason, typically, the number of features that are removed is not very large. Moreover, the procedure of selecting a subset of features cannot usually be considered very effective when solving problems with a very large number of features. Therefore, PCA and LDA algorithms were used to reduce the dimensions of the dataset in the present study. In this method, the data is converted from a complex space with high dimensions into a simple space with low dimensions.

First, the optimal dimensions were obtained through running the PCA algorithm on the data. As a result, the number of data dimensions decreased from 561 to 250. Figure 2 displays the amount of changes for different dimensions of PCA. Based on these results, the optimal number of dimensions was 250, and this number was, thus, considered for the selection of features. With respect to the LDA procedure, it can be maintained that it is similar to PCA. Nevertheless, it should be remembered that the main purpose of LDA analysis is to find the conversion that distinguishes groups to the most extent, whereas the PCA procedure only reduces the problem dimensions and establishes independence between them.

Since the LDA algorithm is one of the reduction methods, it was also applied to the data, as a result of which the number of dimensions was reduced from 561 to 5.

3.3. Proposed Method based on RBF + SVM

In the domain of mathematical modeling, RBF is an artificial neural network that employs radial basis functions as activity functions. The output of this network is a linear combination of radial basis functions for input parameters and neurons. These networks are used in approximation functions, prediction of time series, categorization, and system control. RBF networks are generally composed of an input layer, a hidden layer with a nonlinear RBF activity function, and an output layer. These networks enjoy a high learning speed. Thus, in this proposed method, features were extracted using RBF. Subsequently, the extracted features were sent in the form of a hidden vector to the support vector algorithm with a radial basis neural network kernel to classify the target dataset.

As mentioned above, the RBF network is a three-layer neural network containing a hidden layer. It uses the interesting Cover's theorem to solve complex and nonlinear problems and has many applications in practice. This neural network adopts a different approach to multilayer perceptron and tends to solve complex problems with very simple and interesting approaches. Unlike the MLP approach whereby the synaptic weights of all layers need to be computed, in the RBF network, the input layer is connected directly to the hidden layer without the synaptic weight existing between the two layers. This network's hidden layer neurons function as a nonlinear kernel (Gaussian RBF) and are responsible for mapping data from nonlinear to linear space. Furthermore, an SVM classifier with a fixed core function is not capable of learning complex features but provides good decision levels through maximizing margin, using soft-margin techniques. It should also be stated that the input of this network can be modeled as a vector of real numbers, and its output is a scalar function of the input vector which is calculated as follows (Equation (1)).

$$\varphi(x) = \sum_{i=1} a_i \rho(||x - c_i||) \qquad (1)$$

where N is the number of hidden layer neurons, c_i denotes the center vector of the neuron i, and a_i represents the weight of the neuron i in the linear output neuron. Figure 3 shows the proposed RBF + SVM.

The only difference between this method and the one described immediately below lies in the RBF network. In the latter, CNN is used in place of RBF [17].

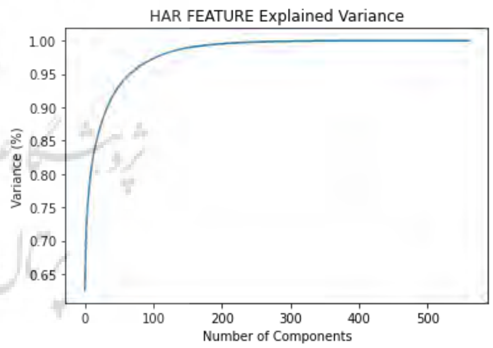

Figure. 2. Number of optimal dimensions in PCA

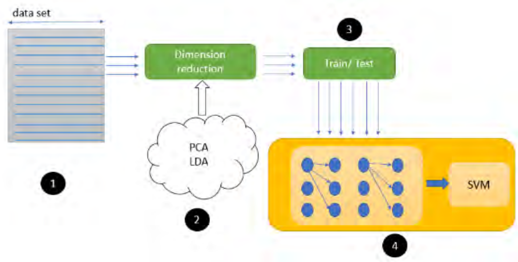

Figure. 3. Outline of proposed method RBF + SVM

3.4. Proposed Method based on CNN + SVM

One of the most important disadvantages of convolutional networks is the existence of fully-connected layers which include the biggest number of learning parameters. These layers are responsible for learning the features which are extracted by the convolutional layers. The last of fully-connected layers are referred to as the score calculator of categories. This layer calculates the outputs of each category making use of a probabilistic function such as Softmax. Next, the error rate of the network is computed by means of a cost function. When it is intended to reduce the amount of network error, the error is propagated into the network, and the weights of each neuron in the fully-connected layers and the amounts of kernels in the convolutional layers are changed through the post-diffusion function. The fully-connected layers in CNN are computationally heavy and time-consuming. Additionally, these layers are less accurate than classifiers such as support vector machines. For this reason, the researchers proposed and made use of a hybrid classifier for the data, which used convolutional networks to extract the features and a support vector machine for classification. The results of the tests were indicative of the superiority of the proposed method over the existing methods in this area [18].

As such, the researchers introduced the Convolution Support Vector Machines (CSVMs) hybrid model, which combines the CNN and SVM classification models. CSVMs are trained by the Stochastic Gradient Descent. This model not only provides high categorization capability for small datasets, but it is also very useful for image data classification and segmentation. Convolutional networks are fully-interconnected MLPs; in these networks, interconnected layers are designed to learn advanced features, and the final layer produces classification results. Convolutional networks are very effective when it comes to learning unchangeable features, but such networks do not always produce the desired classification results. On the contrary, SVM classifiers with a fixed core function are not capable of learning complex features but provide good decision levels by maximizing margin using soft-margin techniques. Regarding these specific characteristics and weaknesses, enumerated above, the researchers focused their attention on producing a hybrid model in which CNN is trained to learn features, and SVM is used for categorization based on features extracted by CNN. With respect to the performance between CNN, MLPs, and SVM, the decision function can be written as follows (Equation (2)).

$$f(x) = (w.\emptyset(x) + b) \qquad (2)$$

where w denotes the weight vector for SVM, b is the bias value, and $\emptyset$ represents all learnable parameters. For SVM, the value of $\emptyset$ is an arbitrary vector. The main purpose of the proposed model, i.e., CSVMs, with respect to the dataset s = {(Xi, Yi)} that $X_i \in R^n$ and $Y_i \in \{+1, -1\}$ it that it must be for a two-class problem (and it can be expanded for even multiple classes). The CSVMs model reduces the following cost function (Equation (3)).

$$Minimize(w) = \left\{\frac{\lambda}{2}||w||^2 + \frac{1}{m}\sum_{i=1}^{m} l(Xi, Yi, w)\right\} \qquad (4)$$

where $l(Xi, Yi; w) = max\ (0,1 - Yif(Xi)\}$ and $\lambda \geq 0$ is the regularization parameter to prevent overfitting, the scale of which is specified by $||w||^2$ To avoid overfitting, the researchers used the random dropout technique. Moreover, $\max(0,1 - Yif(Xi)\}$ represents the Hing-Loss error degrees which is a function for computing error. In a normal case, CNN and SVM models seek to minimize the output of the function $\frac{1}{m}\sum_{i=1}^{m} R(f(Xi), yi)$. but the hybrid model, based on the above relationship, is aimed at optimization [19]. Figure 4 shows the outline of proposed method CNN + SVM.

4. Findings

Most algorithms have parameters which need adjusting during the simulation phase. In this research study, the researchers adjusted the parameters for simulation according to Table 1. The model parameter space was continuous, and this made it possible to obtain different accuracy degrees by changing the parameter values. What follows the tables is an elaboration on some of these parameters.

Batch size: This parameter shows the number of data samples which are distributed through the network. The larger this parameter, the more memory it consumes; by contrast, the smaller its value, the longer the training time. For this purpose, different values (i.e., 8, 16, 32, 64, 128, 256, and 512) were tested. These values are mentioned in [20], which can be useful for batch size.

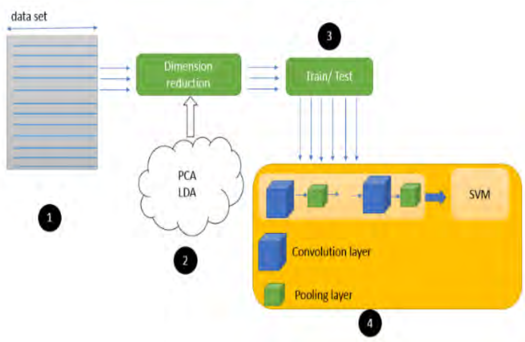

Figure. 4. Outline of proposed method CNN + SVM

Dimensions of the hidden layer: Determining the number of neurons in hidden layers is an important part of the decision regarding the overall architecture of neural networks. In practice, these layers are meant to learn the features extracted by other layers of the network. Using a limited number of neurons in hidden layers results in overfitting. By the same token, using too many neurons in hidden layers can cause several problems, for example, overfitting. Thus, when deciding on the number of neurons in hidden layers, a balance must be struck. With this reason in mind, various values (that is 8, 16, 32, 64, 128, and 256) were taken into consideration. These values are mentioned in [20], which can be useful when determining the size of the hidden layer.

To evaluate and then compare the performance of the proposed method with the basic methods on the UCI HAR dataset, the researchers considered the following criteria: accuracy, precision, recall, and F_score. Table 2 displays the values of these criteria from the dataset as a result of conducting the proposed method. The outcomes of this method were compared with those of several other methods, including support vector machine, decision tree, and random forest. The following present the results of different algorithms on the dataset. The combination of the three algorithms SVM, random forest, and KNN with PCA led the researchers to achieve a maximum accuracy of 96.47, which is the case when using SVM. This reveals that employing SVM can be useful in combining a convolutional network. Additionally, combining different algorithms with PCA + LDA, which are both dimension reduction methods, resulted in a maximum accuracy of 95.69%, which was 0.01% less than using PCA alone.

In line with this method, combining different algorithms with LDA, considered one of the dimension reduction methods, was applied to the data. This caused the number of dimensions to decrease from 561 to 5. Finally, after the three algorithms SVM, random forest, and KNN were combined with LDA, it was observed and reported that SVM reached 96.30% accuracy, which was the highest accuracy degree. The results of different combinations and methods are provided in Table 2. The proposed models were evaluated and then compared with three basic methods and also in three modes: PCA, LDA, and PCA + LDA. In the PCA mode, the SVM basic model obtained the highest accuracy. In the LDA mode, however, the basic models were able to achieve very similar accuracy degrees to one another. Also, in the PCA + LDA combination, the data dimensions were reduced to 5 dimensions, and the KNN model arrived at the highest accuracy.

Of the RBF-based models, the RBF + SVM model obtained the highest accuracy of 99.01%. CNN-based models, too, arrived at acceptable accuracy. Moreover, the CNN algorithm, without considering reduction models, reached an accuracy degree of 98.04%, being the highest one among the CNN models studied. As can be seen from the results displayed in Table 2, reducing the dimensions on the dataset did not lead to much difference when compared to the method whereby the dataset dimensions were not reduced. Figure 5 indicates the accuracy, precision, recall, and F1 diagrams of machine learning models. This figure also shows the RBF and CNN models on the dataset, on which principal component analysis and linear discriminant analysis were performed. Referring to these diagrams, one can observe that the RBF + SVM model reached the highest degree of accuracy.

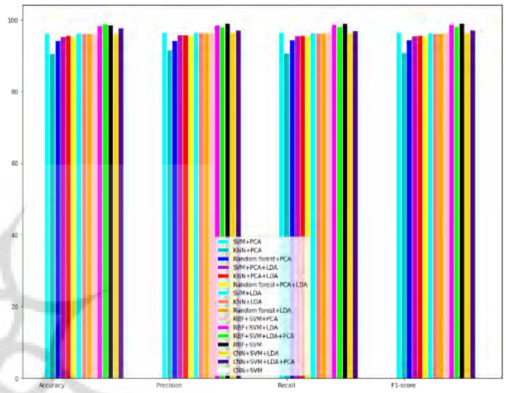

Figure. 5. Evaluation diagrams of machine learning methods and RBF model

Table 1. Details of proposed method on dataset

Parameter	***Value***
Training data volume	70%
Test data volume	30%
Fully-connected layer function	Softmax

Table 2. Values of criteria of proposed method and other algorithms

Model	***Criterion***			
	Recall	***Precision***	***F_score***	***Accuracy***
SVM + PCA	96.37	96.55	96.43	96.47
KNN + PCA	90.52	91.66	90.72	90.94
Random Forest + PCA	94.17	94.27	94.47	94.47
SVM + PCA + LDA	95.39	96.76	95.50	95.52
KNN + PCA + LDA	95.60	95.89	95.69	95.69
Random Forest + PCA + LDA	95.37	95.65	95.46	95.45
SVM + LDA	96.28	96.52	96.35	96.3
KNN + LDA	96.17	96.46	96.25	96.2
Random forest + LDA	96.10	96.39	96.17	96.13
RBF + SVM + PCA	96.12	96.66	96.65	96.65
RBF + SVM + LDA	98.42	98.54	98.76	98.68
RBF + SVM + LDA + PCA	98.91	98.12	98.12	98.12
RBF + SVM	**98.60**	**98.99**	**99.01**	**99.01**
CNN + SVM + PCA	96.20	96.44	96.34	96.36
CNN + SVM + LDA	98.80	97.11	97.02	97.08
CNN + SVM + LDA + PCA	97.13	97.14	97.13	97.48
CNN + SVM	97.42	97.44	97.43	98.04

4.1. Error Analysis

The errors of different RBF and CNN models alongside their accuracy degrees in 40 different execution steps are provided in the Figure 6. The tangentiality of the lines of these error diagrams, as well as the accuracy degrees, indicate that neither overfitting nor underfitting occurred in the proposed methods. The logarithmic shape of the models illustrates this. In [21], various analyses have been offered and considered for both overfitting and underfitting, based on which it can be concluded that the proposed model in the present study solved these issues well.

5. Comparison and Evaluation

The present study proposed two methods. To evaluate these two methods, the researchers used three traditional machine learning approaches, in conjunction with the models presented in a number of studies: [21, 22, 23, 24]. The basis of the method introduced in [7] involved combining the stochastic forest algorithm model and Logistic Model Tree (LMT), which the authors proposed. This method obtained an accuracy degree of 94.02%. In [8], the multi-channel convolutional network was composed of several inputs of different sizes. Each channel could extract the features separately. The extracted features were merged by different filters and channels in the integration layer. The feature map extracted in each layer provided a wide range of features for data classification with a model classification accuracy of 95.25%. In turn, the [9] approach employed a combination of LSTM and CNN. The LSTM network received the time series features which were extracted in spatial feature formats from the convolutional network. In this approach, two LSTM layers were used, and the outputs were sent to two convolutional layers with one Max Pooling layer. As the last step, the categories were normalized. This approach achieved 95.78% accuracy. In [10], GRU was first used to extract the features. Using two gates (i.e., resets and updates), it mapped the extracted features to hidden vectors, which were in fact the same time series properties extracted from the input data. Following this step, the extracted features were given to fully-connected layers, and these layers were used for feature learning. In the multi-input mode, the inputs were fed to several convolutions simultaneously. This approach was also able to achieve 96.20% accuracy.

As stated previously, due to the use of a hybrid structure, the proposed method of this study proved to be more accurate than other approaches. In this proposed method, the RBF network performed data extraction, and the SVM performed the classification better compared to fully-connected layers. In the previous approaches, the main error was the application of fully-connected layers. Therefore, the method adopted in the present study minimized the errors by eliminating these layers. Figure 7 shows the accuracy of the proposed method. In this figure, the proposed method and four other models (LMT [7], multi-channel convolutional neural network [8], CNN-LSTM [9], and multi-input CNN-GRU [10]) are analyzed through comparison.

Diagrams representing machine learning methods, deep learning, and previous research are displayed on the dataset in Figure 5. Likewise, this comparison revealed the fact that the highest accuracy was obtained through the RBF + SVM model.

As displayed in Figure 5, the accuracy of the proposed method was measured at 99.01%. This was indicative of the better performance of the proposed method when compared to other methods. Table 3 shows the difference between the performance of the proposed method and that for each of the other methods.

With respect to the tables provided, it can be seen that the proposed RBF + SVM model obtained the highest accuracy by comparison with all the previous models, including machine learning models. This method first used RBF to extract the features, and the extracted features were then given as input to the SVM model. In order to further compare the proposed method with the other solutions, the researchers took into account the four criteria of accuracy, precision, recall, and F_score. The models were simulated using the Python software and its libraries. The results revealed that the proposed model outperformed all the other four models (i.e., LMT [7], multi-channel convolutional neural network [8], CNN-LSTM [9], and multi-input CNN-GRU [10]) in terms of the above-mentioned four criteria.

6. Conclusion and Suggestions

The principal innovation of this research study was the presentation of a new method for recognizing human activities by means of a combination of support vector machine and radial

Table 3. Accuracy of proposed method compared to other methods

Ref	*Method*	*Accuracy*
Proposed method	RBF + SVM	99.01
[6]	Random forest and logical regression	94.02
[7]	Multi-channel convolutional network	95.25
[8]	Combination of LSTM and CNN	95.78
[9]	Multi input CNN-GRU	96.20

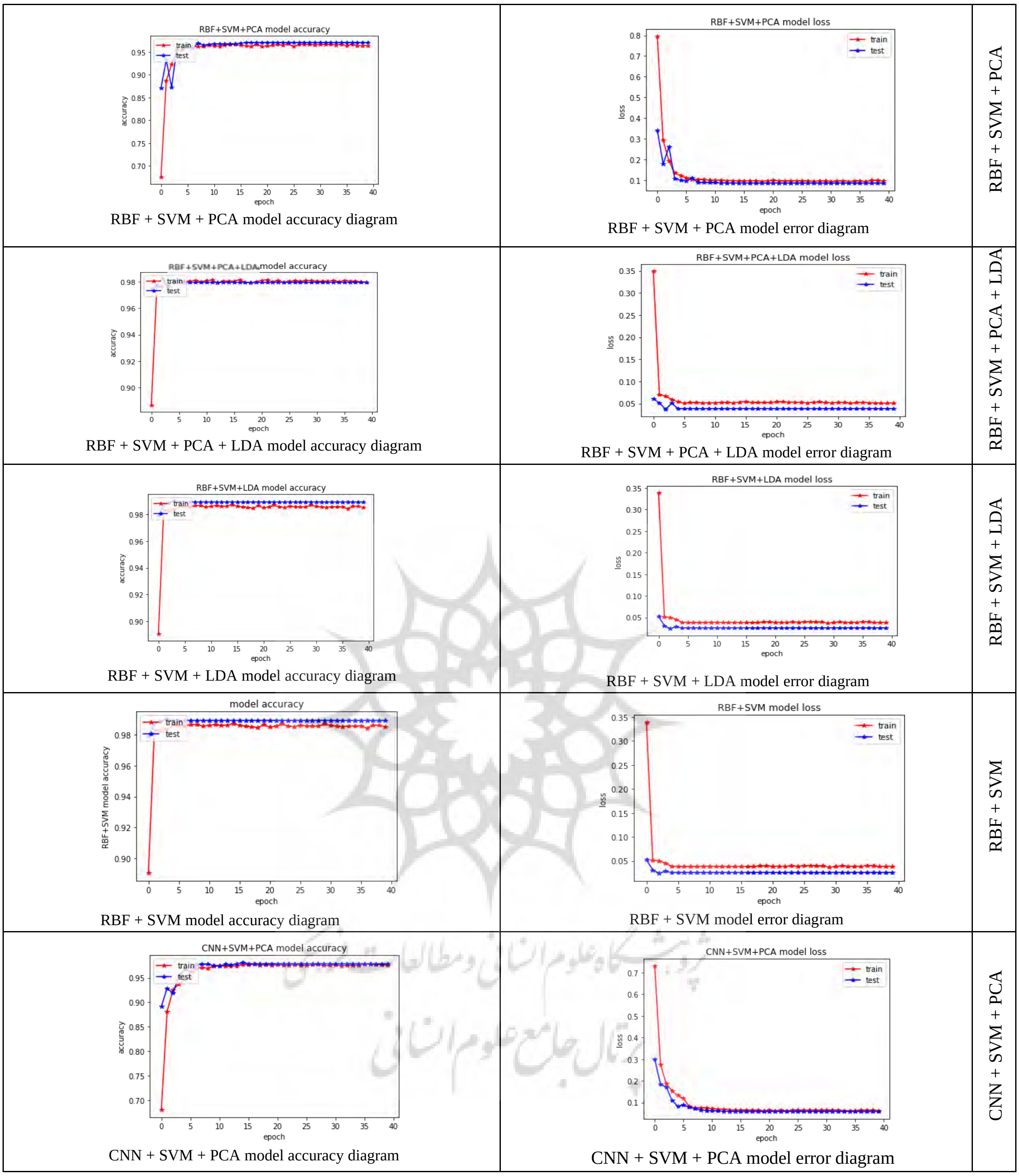

Figure. 6. Accuracy of the proposed methods in different scenarios

basis neural networks. A thorough review of the related literature showed that most of the previous methods have high computational complexity in terms of the number of end-layer neurons. The need for an appropriate method to recognize human activities was, thus, essential to bridge the existing research gap in this regard. Therefore, the present study introduced a model which was intended to have less computational complexity in addition to enhancing accuracy. In the course of this research project, how to use the neural network algorithm to extract the features and how to program the support vector machine to select the features and recognize human activities were proposed as a model.

The findings indicated that using this combination improved the speed of the algorithm. It should be noted that the proposed RBF + SVM model was presented with the aim of extracting features automatically and using the SVM classifier. In this proposed model, the presence of fully-connected

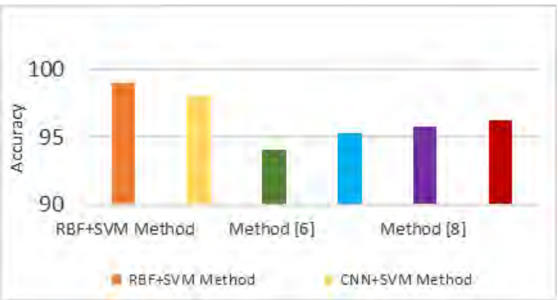

Figure. 7. Comparison of proposed method with other methods in terms of accuracy

layers in RBF led to a decrease in the accuracy of the model and increased its calculations (i.e., added to the computational complexity). To avoid these two problems to some extent, the researchers removed these layers and replaced them with SVM. In order to compare the performance of the proposed method with the other available models, four criteria of accuracy, precision, recall, and F_score were introduced. The methods were simulated by the Python software and its libraries. The output of the comparisons revealed that, with regard to the four criteria, the proposed method had significantly better performance than the four previously-presented models. Put differently, the accuracy, precision, recall, and even F_score of the proposed method proved to be higher than those of the other algorithms. Finally, to prove the claim that the proposed model was better, the researchers considered error diagrams and accuracy diagrams of this model, which showed that there was no overfitting or underfitting. Moreover, based on the results, the reduction of dimensions on the dataset did not lead to much difference compared to the case in which the dataset dimensions were not reduced. More specifically, the results of the present study revealed that the proposed method is 99.01% more efficient than the other algorithms and methods, and it can be used to recognize human activities in different dimensions.

For future research in this area, it is suggested that the proposed method be applied to a larger dataset. Furthermore, capsule neural networks can be used because in many classification problems, they have achieved much better results than conventional neural networks, as well as convolutional networks. Additionally, optimization algorithms such as particle swarm optimization (PSO) algorithms can be used to adjust the basic parameters of the model or to combine deep learning and traditional learning models such as XGBoost.

Declarations

Funding

This research did not receive any grant from funding agencies in the public, commercial, or non-profit sectors.

Authors' contributions

FA: Study design, acquisition of data, interpretation of the results, statistical analysis, drafting the manuscript;

AG: interpretation of the results, revision of the manuscript.

Conflict of interest

The authors declare that no conflicts of interest exist.

References

[1] S. Khalifa, G. Lan, M. Hassan, A. Seneviratne, and S. K. Das, "Harke: Human activity recognition from kinetic energy harvesting data in wearable devices," *IEEE Transactions on Mobile Computing*, vol. 17, no. 6, pp. 1353-1368, 2017. https://doi.org/10.1109/TMC.2017.2761744

[2] M. J. Shayegan" ,A bibliometric investigation into the literature of semantic reasoning in internet of things," *Internet Technology Letters*, vol. 6, no. 3, p. e401, 2023. https://doi.org/10.1002/itl2.401

[3] P. Wang, W. Li, P. Ogunbona, J. Wan, and S. Escalera, "RGB-D-based human motion recognition with deep learning: A survey," *Computer Vision and Image Understanding*, vol. 171, pp. 118-139, 2018. https://doi.org/10.1016/j.cviu.2018.04.007

[4] A. Yurtman and B. Barshan, "Human activity recognition using tag-based radio frequency localization," *Applied Artificial Intelligence*, vol. 30, no. 2, pp. 153-179, 2016. https://doi.org/10.1080/08839514.2016.1138787

[5] S. Seneviratne *et al.*, "A survey of wearable devices and challenges," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 4, pp. 2573-2620, 2017. https://doi.org/10.1109/COMST.2017.2731979

[6] J. Wang, Y. Chen, S. Hao, X. Peng, and L. Hu, "Deep learning for sensor-based activity recognition: A survey," *Pattern recognition letters*, vol. 119, pp. 3-11, 2019. https://doi.org/10.1016/j.patrec.2018.02.010

[7] H. Nematallah, S. Rajan, and A.-M. Cretu, "Logistic model tree for human activity recognition using smartphone-based inertial sensors," in *2019 IEEE SENSORS*, Montreal, QC, Canada, IEEE, 2019, pp. 1-4. https://doi.org/10.1109/SENSORS43011.2019.8956951

[8] [^] N. Sikder, M. S. Chowdhury, A. S. M. Arif, and A.-A. Nahid, "Human activity recognition using multichannel convolutional neural network," in *2019 5th International conference on advances in electrical engineering (ICAEE)*, Dhaka, Bangladesh, IEEE, 2019, pp. 560-565. https://doi.org/10.1109/ICAEE48663.2019.8975649

[9] K. Xia, J. Huang, and H. Wang, "LSTM-CNN architecture for human activity recognition," *IEEE Access*, vol. 8, pp. 56855-56866, 2020. https://doi.org/10.1109/ACCESS.2020.2982225

[10] N. Dua, S. N. Singh, and V. B. Semwal, "Multi-input CNN-GRU based human activity recognition using wearable sensors," *Computing*, vol. 103, no. 7, pp. 1461-1478, 2021. https://doi.org/10.1007/s00607-021-00928-8

[11] Y. Li, G. Yang, Z. Su, S. Li, and Y. Wang, "Human activity recognition based on multienvironment sensor data," *Information Fusion*, vol. 91, pp. 47-63, 2023. https://doi.org/10.1016/j.inffus.2022.10.015

[12] S. K. Challa, A. Kumar, and V. B. Semwal, "A multibranch CNN-BiLSTM model for human activity recognition using wearable sensor data," *The Visual Computer*, vol. 38, no. 12, pp. 4095-4109, 2022. https://doi.org/10.1007/s00371-021-02283-3

[13] Y. Kaya and E. K. Topuz, "Human activity recognition from multiple sensors data using deep CNNs," *Multimedia Tools and Applications*, vol. 83, no. 4, pp. 10815-10838, 2024. https://doi.org/10.1007/s11042-023-15830-y

[14] F. S. Abuhoureyah, Y. C. Wong, and A. S. B. M. Isira, "WiFi-based human activity recognition through wall using deep learning," *Engineering Applications of Artificial Intelligence*, vol. 127, p. 107171, 2024. https://doi.org/10.1016/j.engappai.2023.107171

[15] X. Chen, Y. Zou, C. Li, and W. Xiao, "A deep learning based lightweight human activity recognition system using reconstructed WiFi CSI," *IEEE Transactions on Human-Machine Systems*, vol. 54, no. 1, pp. 68-78, 2024. https://doi.org/10.1109/THMS.2023.3348694

[16] H. Feng, Q. Shen, R. Song, L. Shi, and H. Xu, "ATFA: Adversarial Time–Frequency Attention network for sensor-based multimodal human activity recognition," *Expert Systems with Applications*, vol. 236, p. 121296, 2024. https://doi.org/10.1016/j.eswa.2023.121296

[17] Y. Wu, H. Wang, B .Zhang, and K.-L. Du, "Using radial basis function networks for function approximation and classification," *International Scholarly Research Notices*, vol. 2012, 2012. https://doi.org/10.5402/2012/324194

[18] S. Ji, W. Xu, M. Yang, and K. Yu, "3D convolutional neural networks for human action recognition," *IEEE transactions on pattern analysis and machine intelligence*, vol. 35, no. 1, pp. 221-231, 2012. https://doi.org/10.1109/TPAMI.2012.59

[19] K. Simonyan and A. Zisserman, "Two-stream convolutional networks for action recognition in videos," *Advances in neural information processing systems*, vol. 27, 2014, pp. 568-576. https://dl.acm.org/doi/10.5555/2968826.2968890

[20] R. Mousa, S. Nazari, A. K. Abadi, R. Shoukhcheshm, M. N. Pirzadeh, and L. Safari, "TI-Capsule: Capsule Network for Stock Exchange Prediction," *arXiv preprint arXiv:2102.07718*, 2021.

[21] https://doi.org/10.48550/arXiv.2102.07718

[22] H. Zhang, L. Zhang, and Y. Jiang" ,Overfitting and underfitting analysis for deep learning based end-to-end communication systems," in *2019 11th International Conference on Wireless Communications and Signal Processing (WCSP)*, Xi'an, China, IEEE, 2019 pp. 1-6 . https://doi.org/10.1109/WCSP.2019.8927876

Fatemeh Azimzadeh is a researcher currently serving as an Assistant Professor at the Academic Center for Education, Culture, and Research (ACECR) in Tehran, Iran. She holds a Bachelor's degree from Tehran University and has furthered her education with a Master's and PhD from Universiti Putra Malaysia (UPM). Dr. Azimzadeh has made her contributions to the field of information technology, particularly in the areas of information retrieval, web search and NLP. Currently, she is the Head of the Scientific Information Database (SID) at ACECR. Her vision is enhancing the accessibility and quality of scientific resources available to researchers and scholars.

Dr. Ali Ghoroghi is a Research Fellow at Cardiff University's Urban Intelligence Research Centre and a Senior AI Engineer at Optimise-AI. His advanced work in AI and machine learning has resulted in over 15 peer-reviewed publications and leadership roles in significant European and UK-funded projects.