Automatika
Journal for Control, Measurement, Electronics, Computing and Communications

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/taut20

HARNet: automatic recognition of human activity from mobile health data using CNN and transfer learning of LSTM with SVM

R. Anandha Praba & L. Suganthi

To cite this article: R. Anandha Praba & L. Suganthi (2024) HARNet: automatic recognition of human activity from mobile health data using CNN and transfer learning of LSTM with SVM, Automatika, 65:1, 167-178, DOI: 10.1080/00051144.2023.2290736

To link to this article: https://doi.org/10.1080/00051144.2023.2290736

AUTOMATIKA
2024, VOL. 65, NO. 1, 167–178
https://doi.org/10.1080/00051144.2023.2290736

HARNet: automatic recognition of human activity from mobile health data using CNN and transfer learning of LSTM with SVM

R. Anandha Praba[a] and L. Suganthi[b]

[a]Department of ECE, Meenakshi College of Engineering, Chennai, India; [b]Department of BME, SSN College of Engineering, Chennai, India

ABSTRACT
Human Activity Recognition (HAR) system is analysing human behaviour using mobile health technology. Mobile Health data (MHEALTH) uses electronic devices to collect data and identify the activity of the patient in real-time. Recordings of 10 patients' vital signs from various circumstances are included in the dataset. With a sensor attached to their bodies, they were required to carry out a number of physical tasks. Due to the lack of accuracy in the other state-of-the-art algorithms, we proposed Human Activity Recognition Neural Network (HARNet) architecture for automatic recognition of human activity using CNN and LSTM with the transfer learning of SVM. Here, the human health behaviour was analysed and classified using different ML and DL algorithms. The hybrid techniques of CNN and LSTM are selected across the different DL algorithms and it is used to extract independent and discriminating features, which aids the SVM classifier to attain good classification. When compared to other DL methods, HARNet performed better, achieving 99.8% accuracy. Overall, HAR systems have many potential applications in various fields, including healthcare, wellness, sports and surveillance. They have relationships to many different academic disciplines, including sociology, human–computer interaction, medical and may offer individualized help for different domains.

ARTICLE HISTORY
Received 8 July 2023
Accepted 14 November 2023

KEYWORDS
Human Activity Recognition; Mobile Health data (MHEALTH); Convolutional Neural Network (CNN); Long Short Term Memory (LSTM); Support Vector machine (SVM); deep learning

1. Introduction

Due to the accessibility of sensors and accelerometers, the field of HAR has emerged as one of the most popular research fields [1,2]. To analyse the different human behaviours connected to health and physical activities, such as standing, lying down, sitting and walking mobile health data can be employed. The MHEALTH dataset is a valuable resource for researchers and healthcare providers interested in mobile health human behaviour analysis. It can be accessed through various platforms, including Kaggle, MHEALTH Research Group and Research Gate. Additionally, the use of sensing capabilities in MHEALTH data collection applications and MHEALTH apps can enhance the collection and analysis of health-related data.

In order to evaluate changes in the likelihood of developing chronic diseases, mobile sensing data can be utilized to estimate levels of stress indicators [3]. By validating it against measures of actual and perceived mental health, stress, sleep length and inflammation, a mobile sensing collection instrument can be utilized to measure stress levels [4]. Investigating differences in health attitudes and behaviours among users of mobile health apps can be done with their help. After adjusting for variables in mobile health it will enhance the chance of health management behaviours [5,6].

In their investigation of human behaviour with regard to mobile health, Hilty DM and Chan S looked at smartphone/devices, apps and cognition. Overall, Mobile Health Human Behaviour Analysis can provide insights into human behaviour related to health management and physical activities using mobile health technology [7].

Machine learning has various applications in MHEALTH data analysis, including classification and prediction [8], feasibility of machine learning techniques [9], predictive analytics [10], clinical decision support, population health management, deep learning and collaboration among machine learning researchers. These applications can help in the development of effective interventions and treatments for various health conditions.

Data gathered from several sensors implanted in various locations on the human body is used to identify the human activity using a model based on DL. Here, the hybrid CNN and LSTM with the transfer learning of SVM architecture is designed to perform the Human Activity Recognition (HAR). The proposed model produces the better accuracy compare to other state of art algorithms.

This study demonstrates the HAR system using HARNet architecture. Section 2 reviews the existing ML

CONTACT R. Anandha Praba ranandhapraba@gmail.com Department of ECE, Meenakshi College of Engineering, Chennai, Tamilnadu, India

and DL for recognition of human activity based on their performance. Section 3 describes the proposed HAR-Net model used in this work. The performance and the visualization of the data are further described in Section 4. Section 5 brings the paper to a close and identifies potential areas for further research.

2. Related work

Recognizing human behaviour has proven useful for helping clinicians treat and remotely monitor patients. It is crucial for medical treatment and diagnosis as well as determining the likelihood that a patient may become ill or pass away from specific illnesses or health issues [11].

Machine learning can be used to analyse human activity data in various ways. The machine learning can provide insights into mobile health human behaviour data using various methods. HAR using DL models has become increasingly popular due to its high effectiveness in recognizing complex tasks. The deep learning algorithms such as CNN, RNN, Deep Belief Network, XGBoost and Auto encoder have been used for HAR with wearable and mobile sensor networks. When compared to more conventional machine learning methods, these algorithms have demonstrated encouraging results in the recognition of complicated tasks while also being very inexpensive.

The HAR health Model is the best way to safeguard patients who are at risk of developing chronic heart failure, other ailments, or contracting a disease. It can also assist with remote monitoring of patients and provide an immediate response when a fall is detected [12].

For the purpose of producing incredibly precise performance results, Mo et al. [13] concentrate on feature extraction and data pre-processing. Here, the author hybrid the Multilayer perceptron with the CNN. In this hybrid method, the CNN is for feature extraction. The MLP is used for activity classification and it achieves 81.8% accuracy which is higher compared to other state-of-the-art algorithms. The author demonstrates the hybrid method using CAD-60 Dataset.

Islam et al. [14] suggested using a classification model to predict myocardial disease. In this study, the authors attempt to predict human fitness by gathering information on activity levels at specific times, such as running, cycling and standing, as well as blood pressure readings. They employed the clustering model, and their accuracy was 84.66%.

R. Al-Ghannam et al [15] created a mobile-based application that uses accelerometer data from mobile phones that were collected during six prayers for a total of 118 samples to identify different types of prayer movements, such as standing, bowing, prostrating and sitting. The accuracy of prayer activity recognition is performed using a variety of popular classifiers, such as J48 Decision Tree, Naive Bayes. The comparison of the algorithms is explained using data mining tool kit (WEKA). Their respective categorization accuracy rates are 91.3%, 90.2% and 88%.

Yu et al. [16] use block chain technology to benefit the sports sector by gathering health data. According to the publication, block chain is suitable for the sports business since it is decentralized and safe.

In order to forecast human fitness, Kerdjidj et al. [17] detect the different daily activities Living using different machine learning algorithms. After combining the readings from the accelerometer and gyroscope, the data collecting and classification task initially only use the accelerometer's data.

The recruitment processes for javelin throwing were improved utilizing neural networks, which were based on a mathematical model developed in this study and attempted generalization. The Neural Network is used to employ the recruiting and selection process for javelin throwing [18].

In [19], Convolutional neural networks, a very powerful deep learning technology that can accurately represent features, are used by the author to demonstrate HAR. Triaxial sensors and Inertial Measurement Unit sensors are taken into consideration as input devices and placed six sensors: two on the left and right shanks, two in the middle of the foot and one on the lumbar region.

In [1], the author explains a data-collecting module prototype that they created, which collects patient data and identifies problematic health statuses to enable early intervention. Here, an accelerometer built inside a smartwatch is employed as an input device for arm position detection, and filtering, normalization, and feature extraction are used as pre-processing techniques. On chest is taken into consideration as input device for body posture recognition.

2.1. Main contribution and challenges

- Multiple people: It can be challenging to map the actions of multiple residents when sensors are installed in the home environment because there may be multiple occupants there.
- During data collection: If the data are to be acquired via sensors, the user must wear many sensors, and the location of the sensors is problematic since it influences the outcomes.
- Multiple tasks: It is challenging to identify someone who is performing multiple tasks at once.
- Classification models with overfitting and underfitting potential include decision trees and neural networks.
- Sensor limitations: When using malfunctioning sensors, we are unsure whether the data are accurate.
- Real-time data: A lot of results that were calculated using conventional datasets could change when real-time datasets are used.

- When there are insufficient training data, underfitting can occur when using a hybrid CNN with LSTM and SVM. Therefore, the implementation strategy must be in line with the data.

3. Proposed methodology

The three key steps in the creation of the proposed system, known as HARNet, are data collection from Kaggle, Selection of deep learning algorithm and perform transfer learning. This entire processing was carried out in Python. Figure 1 shows the architecture of HARNet system.

The pre-trained model is trained with large volume of data and it can be stored for further process. It can be used as a starting point for a new machine-learning task, instead of building a model from scratch. Pre-trained models can be identified by their architecture and the dataset they were trained on. Some examples of pre-trained models are Convolutional Neural Networks, VGG [20], ResNet [20] and Inception. A pre-trained model can be applied to a task either directly or after being modified through transfer learning. Transfer learning includes taking advantage of the representations that a prior network has learnt to identify significant features in fresh samples. The pre-trained model can then be combined with a fresh classifier that is being trained from start so that the previously learnt feature mappings can be applied to the new dataset.

3.1. The contributions of this article are highlighted as below

(1) Data Collection: Collect a dataset of labelled data. The data should be labelled with the different human behaviours that you want to classify.
(2) Pre-processing: The data can be pre-processed by cleaning and filtering the sensor data collected from mobile and wearable devices.
(3) Feature extraction: The CNN layers are used for feature extraction on the pre-processed data. The CNN layers can excerpt spatial features.
(4) Sequence modelling: The LSTM layers are used for sequence modelling on the extracted features by the CNN layers.
(5) Training LSTM Model: Train an LSTM network on the labelled data. The LSTM network could learn the features of the data that are associated with each human behaviour.
(6) Fine Tuning: Once the LSTM network is constructed, then freeze the weights of the network. This will prevent the weights from being updated during the next step.
(7) Transfer Learning with SVM: Add a new layer to the LSTM network. The new layer should be a SVM classifier.
(8) Training SVM: Train the Support Vector Machine classifier on the labelled data. The SVM classifier should be trained on the extracted features by the LSTM network.

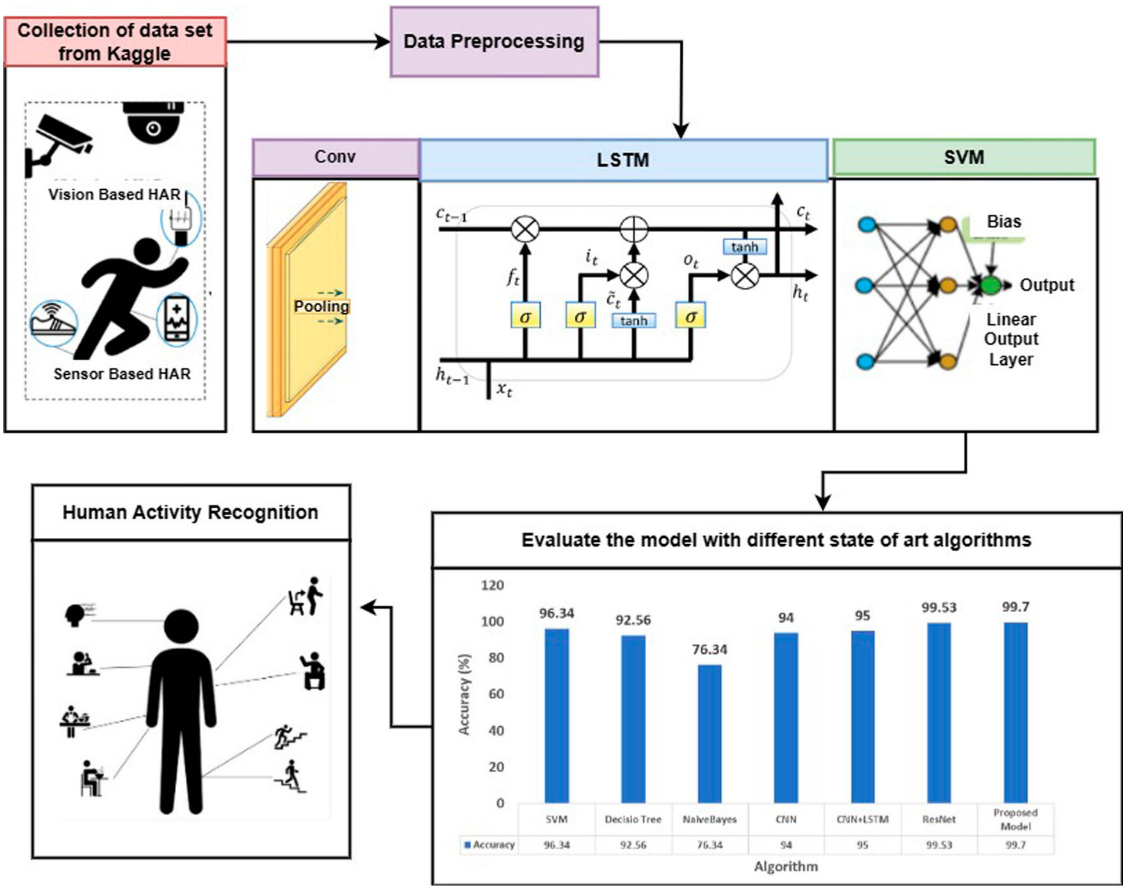

Figure 1. Human Activity Recognition System Using HARNet.

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1215745 entries, 0 to 1215744
Data columns (total 14 columns):
 #   Column    Non-Null Count    Dtype
---  ------    --------------    -----
 0   alx       1215745 non-null  float64
 1   aly       1215745 non-null  float64
 2   alz       1215745 non-null  float64
 3   glx       1215745 non-null  float64
 4   gly       1215745 non-null  float64
 5   glz       1215745 non-null  float64
 6   arx       1215745 non-null  float64
 7   ary       1215745 non-null  float64
 8   arz       1215745 non-null  float64
 9   grx       1215745 non-null  float64
 10  gry       1215745 non-null  float64
 11  grz       1215745 non-null  float64
 12  Activity  1215745 non-null  int64
 13  subject   1215745 non-null  object
dtypes: float64(12), int64(1), object(1)
memory usage: 129.9+ MB
None
```

Figure 2. Dataset visualization.

(9) Classification: Once the SVM classifier is trained, it can use it to classify new data. The new data should be processed by the LSTM network and then passed to the SVM classifier for classification. Here, Adam optimizer is used to build the model.

3.2. Data collection and pre-processing

HAR is working based on supervised or unsupervised ML Algorithm. The unsupervised ML does not require the label in the data set. Based on the distance between the data points, it groups the data but the supervised HAR system requires prior training with a labelled data set. In the proposed work, the MHEALTH Dataset is used, which is available at https://www.kaggle.com/datasets/nirmalsankalana/mhealth-dataset-data-set-csv.

The body movements and vital signs of 10 participants while they engaged in a variety of physical activities are captured in the MHEALTH (Mobile HEALTH) dataset as shown in figure 2. Sensors that keep track of the acceleration, rate of turn and magnetic field orientation of various body parts are mounted on the subject's chest, right wrist and left ankle. Shimmer2 [BUR10] wearable sensors were used for the recordings [11].

The sensor positioned on the chest also provides 2-lead ECG measurements which are not used for the development of the recognition model but rather collected for future work purposes. This information can be used, for example, for basic heart monitoring, checking for various arrhythmias or looking at the effects of exercise on the ECG. All sensing modalities are recorded at a sampling rate of 50 Hz, which is considered sufficient for capturing human activity. The supervised HAR system in the HARNet architecture is intended to identify the following 12 human activities [21,22].

Class Label 01: Standing still (1 min)

Class Label 02: Sitting and relaxing (1 min)

Class Label 03: Lying down (1 min)

Class Label 04: Walking (1 min)

Class Label 05: Climbing stairs (1 min)

Class Label 06: Waist bends forward (20x)

Class Label 07: Frontal elevation of arms (20x)

Class Label 08: Knees bending (crouching) (20x)

Class Label 09: Cycling (1 min)

Class Label 10: Jogging (1 min)

Class Label 11: Running (1 min)

Class Label 12: Jump front & back (20x)

3.3. Hybrid CNN – LSTM

It is a hybrid model that aims to combine the strengths of LSTM and CNNs. In essence, CNN is involved in feature extraction, whereas LSTM is in control of the learning objective. Here, the max-pooling layer is used to introduce the information to the LSTM. After this procedure, one can use a flattened layer to arrange the data in a way that allows a dense (fully connected) layer to read it [20]. The Hybrid CNN-LSTM is represented in Figure 3.

3.4. Convolutional Neural Network

In numerous applications, including image classification, speech recognition, computer vision, object detection, bioinformatics and video analysis, CNN networks have been widely used because of their high performance [23].

The convolutional layer of the CNN architecture is described as follows.

$$\begin{aligned} \mathrm{F}(f,g) &= (M*N)(f,g) \\ &= \sum\sum M(f+r,\ g+t)N(r,t) \end{aligned} \tag{1}$$

Here, the convolutional layer representation is with $(M*N)$. N refers the 2D filter of size $f \times g$, M refers the Input matrix and F stands for the output of a 2D feature map.

The CNN used in HARNET consists of input layer, convolutional layer, batch normalization and fully connected/dense layer. The input layer takes in the raw sensor data. In convolutional layer a set of filters is applied to the input data to extract spatial features. Each filter slides across the input, performing element-wise

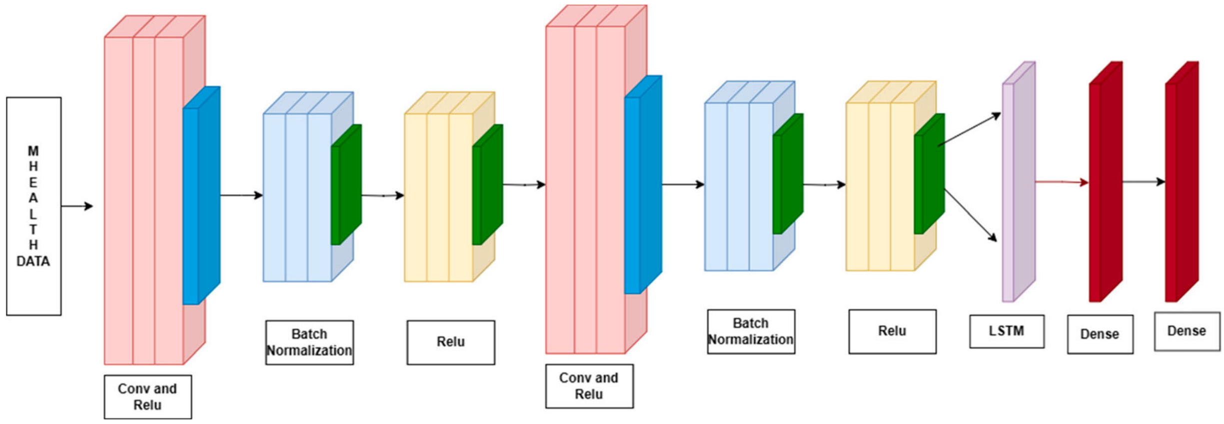

Figure 3. CNN-LSTM architecture used in HARNet.

multiplications and summations. Here activation function (ReLU) is applied after each convolutional operation, Rectified Linear Unit (ReLU) helps introduce non-linearity, allowing the network to learn more complex relationships. The pooling layers down-sample the spatial dimensions, reducing computation and focusing on the most informative features. Batch normalization Helps stabilize and accelerate training by normalizing the input to a layer. The final dense layers connect every neurone in one layer to every neurone in the next layer. The number of neurones in the last dense layer corresponds to the number of activity classes. The activation softmax is applied to get class probabilities.

3.5. Long Short Term Memory

It is capable of remembering information over long periods of time and can selectively forget or remember specific pieces of information. A critical property from the beginning of the input sequence can be transmitted across a great distance via LSTM units, capturing potential long-distance dependencies. The issue of the disappearing gradients is one of the major difficulties in training the standard RNN, though. Small gradient updates hurt the RNN, especially in the earlier layers. As a result, it cannot retain the information for lengthy sequences. Internal loops are a product of LSTM, which keeps only relevant input and discards irrelevant information.

LSTM models are well-suited for HAR systems due to their ability to learn and remember long-term dependencies in sequential data, crucial for accurately recognizing human activities. They can directly learn from raw sensor data, eliminating the need for manual feature engineering, and handling complex temporal patterns in the data. Their recurrent structure allows them to remember previous inputs and use them to inform predictions for future outputs, essential for accurately recognizing human activities involving sequences of complex movements. LSTM models have achieved state-of-the-art results on various HAR datasets and are widely used in commercial HAR products, such as fitness trackers and smartwatches.

For each LSTM cell, three separate gates – the input gate, forget gate and output gate – are used to control the flow of information.

3.5.1. Step 1: The forget gate

As stated in the following formula, the output value of the forget gate is calculated using the input value of the present time and the output value of the most recent moment.

$$f_t = \sigma(We_f.[he_{t-1}, in_t]) + bi_f$$

where the range of values is (0, 1), We_f is the forget gate's weight, bi_f is its bias, in_t is the input value for the present moment, and he_{t-1} is its output value for the previous instant.

3.5.2. Step 2: Learn gate

The sigmoid layer and the tanh layer are two further layers in this process. Using the values 0 and 1, the sigmoid layer must determine whether new data should be updated or discarded. The tanh layer is in charge of passing values between −1 and 1, which allows the weights to be updated. Values are chosen based on their level of significance.

$$i_t = \sigma(We_i[he_{t-1}, in_t] + bi_t)$$

$$C_t = \tanh(we_c[he_{t-1}, in_t] + bi_c)$$

Here, We_i refers the weight of the input gate, We_c stands for the weight of the candidate input gate, bi_c mention the bias of the candidate input gate and bi_t refers the bias of the input gate.

3.5.3. Step 3: The remember gate

The cell contains information that must be remembered at this point. Current cell state is updated using the following formula and the value range of C_t is (0,1).

$$C_t = f_t * C_{t-1} + i_t * C_t$$

```
Model: "model_1"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 input_5 (InputLayer)        [(None, 100, 12)]         0

 conv1d_7 (Conv1D)           (None, 100, 32)           1184

 batch_normalization_7 (Batc  (None, 100, 32)          128
 hNormalization)

 re_lu_6 (ReLU)              (None, 100, 32)           0

 conv1d_8 (Conv1D)           (None, 100, 64)           6208

 batch_normalization_8 (Batc  (None, 100, 64)          256
 hNormalization)

 re_lu_7 (ReLU)              (None, 100, 64)           0

 max_pooling1d_3 (MaxPooling  (None, 50, 64)           0
 1D)

 lstm_3 (LSTM)               (None, 64)                33024

 dense_6 (Dense)             (None, 128)               8320

 dense_7 (Dense)             (None, 13)                1677

=================================================================
Total params: 50,797
Trainable params: 50,605
Non-trainable params: 192
```

Figure 4. Model used in HARNet Architecture.

3.5.4. Step 4: The output gate

The input in_t and the output he_{t-1} are received as input values of the output gate at time t, and the output o_t of the output gate is obtained as follows:

$$o_t = \sigma(We_o[he_{t-1}, in_t] + bi_o)$$

$$h_t = o_t * \tanh(C_t)$$

Figure 4 depicts the model description used in the HARNet using Python environment.

3.6. SVM classifier

To apply SVM as transfer learning, the standard SVM cost function to build a single class object classifier by learning from similar domains is modified. One of the supervised learning models is the SVM, which, given a collection of labelled training data, can classify fresh data points into various groups [25]. The learning paradigm known as "transfer learning" allows us to apply the knowledge we have learned in a particular field to other, more familiar fields. Here, the SVM is used as a transfer learning of LSTM with different modifying the standard SVM cost function. Another transfer learning architecture based on SVM involves a multi-source fast transfer learning algorithm. This algorithm uses MHEALTH data to train the model and it improves the accuracy of an algorithms.

In addition, LSTM-SVM is used with transfer learning for automatic defect classification. Figure 5 shows the LSTM deep learning algorithm with the SVM classifier. Here, the SVM classifier is used as a transfer learning of last layer included in the LSTM [8,25–27].

The mathematical formula for transfer learning with SVM kernel function depends on the specific kernel function being used [28,29]. Here are some examples of SVM kernel functions and their equations:

3.6.1. Linear kernel

When the data can be separated linearly, the linear kernel, which is the simplest kernel function, is employed. The linear kernel's equation is:

$$K(x, y) = x * y$$

Here, $K(x, y)$ is a kernel function and it can be calculated using the multiplication of x and y inputs.

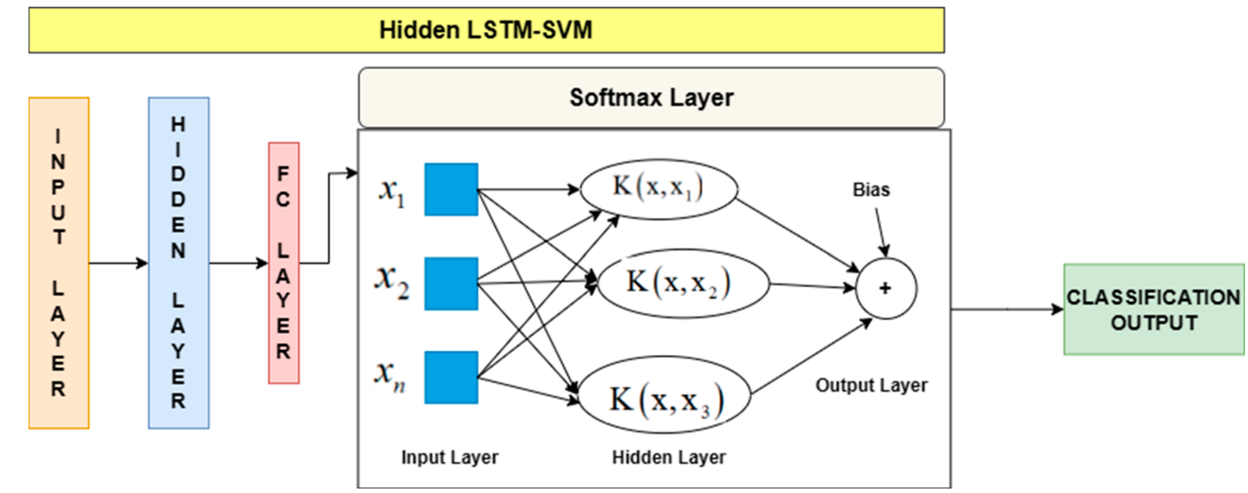

Figure 5. Transfer learning of LSTM with SVM used in HARNet Architecture.

3.6.2. Gaussian (RBF) kernel

A well-liked kernel function that can handle data that cannot be separated linearly is the Gaussian kernel. The Gaussian kernel equation is as follows:

$$K(x, y) = exp(-gamma * ||x - y||^2)$$

where gamma is a parameter that controls the width of the Gaussian.

Here, the Gaussian (RBF) kernel function is working good compare to other kernel function. A labelled data set is used to train the LSTM model initially. The softmax layer in the LSTM is transferred as a SVM kernel function.

Backpropagation is used to minimize the loss function. After that, new data can be classified using the trained SVM model. Here is an illustration of how transfer learning can be used to enhance an SVM model's performance on a new dataset. LSTM transfer learning is used to improve the performance of the SVM model after initially training a CNN using LSTM model on MHEALTH data.

The CNN-LSTM model [30] will learn to extract features that are relevant to the classification task. The output of the LSTM model will then be used to train the SVM model. The SVM model will then be able to classify the label with better accuracy than if it was trained on the new dataset alone.

3.7. Adaptive moment estimation optimizer

This optimizer is used to train the model. Stochastic Gradient Descent, sometimes known as SGD, is an iterative machine learning technique that, after selecting a random weight vector, optimizes the gradient descent throughout each search. The advantages of this are combined with Adaptive Moment Estimation (Adam) [31,32].

(1) Initialize the $\Omega = 0.00001$ for the step size
(2) Estimates of stochastic moment can be identified by $\mho_1, \mho_2 \in [0.0, 1.5]$
(3) The stochastic objective function with parameters β is $\varphi(\beta)$
(4) β_0 the default parameter is initialized
(5) At time stamp t as 0, the first order F_0 and second order S_0 is initialized
(6) While (t = t + 1) until ${}_t$ is not satisfied
    (a) $f_t = \nabla_\beta \varphi_t(\beta_{t-1})$ = gradients at time t
    (b) $F_t = \mho_1 F_{t-1} + (1 - \mho_1) f_t{}^1$ = First order moment
    (c) $S_t = \mho_2 S_{t-1} + (1 - \mho_2) f_t{}^2$ – Second order moment
(7) Update parameter $\beta_t = \beta_{t-1} - \chi * F_t / (\sqrt{S_t} + \lambda)$
(8) End
(9) Returnβ_t

The significance of Adam Optimizer is that it adapts the learning rates for each parameter individually. This means that it can automatically adjust the step size for each parameter based on its importance in the optimization process. This is particularly useful for tasks with high-dimensional and noisy data like HAR. Adam applies bias correction during the first few iterations, which helps stabilize the optimization process. The adaptability to noisy data and efficient convergence are crucial for achieving high performance. The proposed HARNet architecture achieves 99.8% accuracy with the Adam optimizer.

4. Experimental evaluation

4.1. Data collection and pre-processing

The visualization of the many activities in the MHEALTH dataset utilized in the HARNet Architecture is shown in Figure 6. Every session was captured on camera for documentation. This dataset is found to generalize to ordinary daily life activities given the range of body parts participating in each activity (e.g.

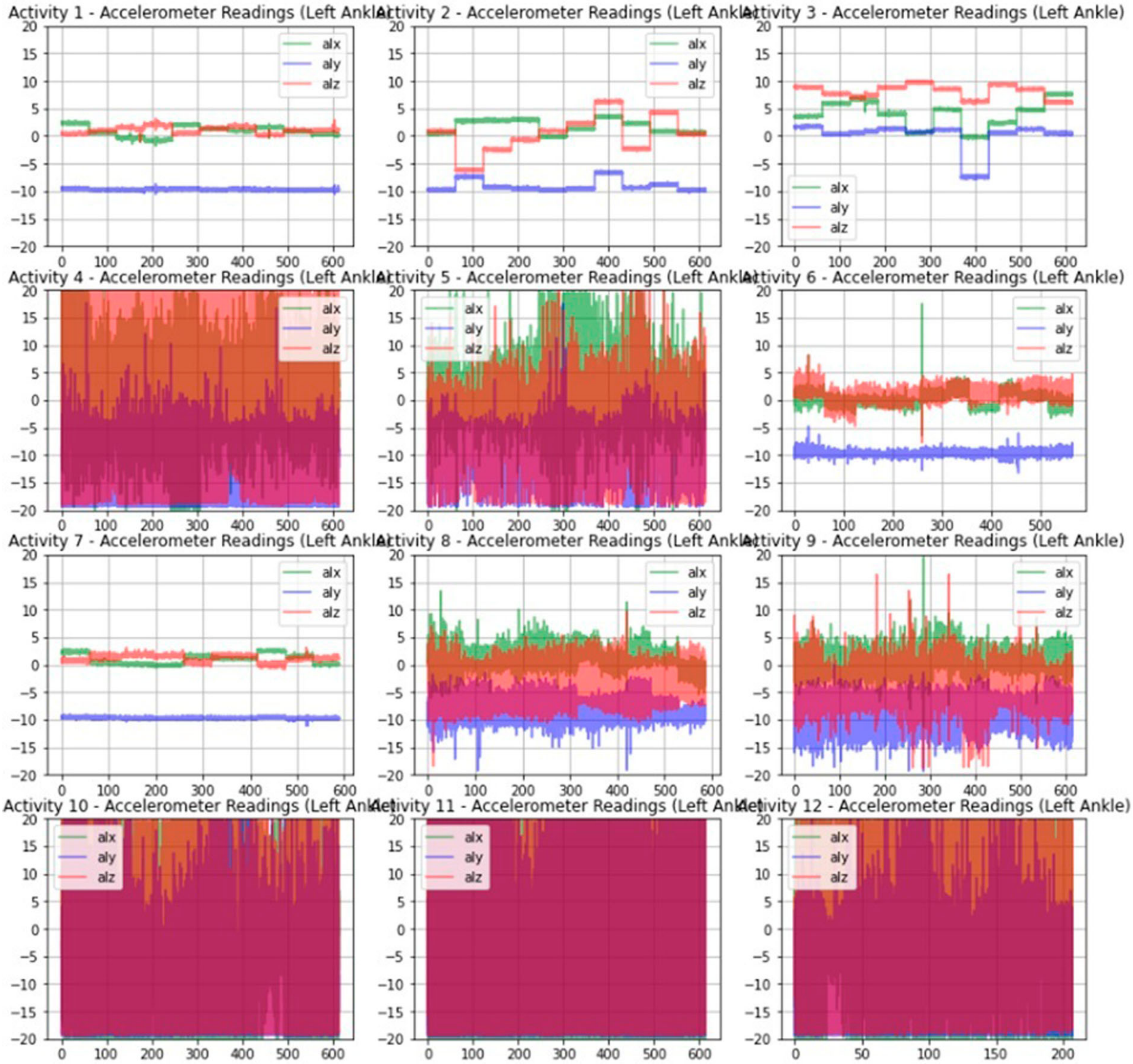

Figure 6. Visualization of the list of activities in MHEALTH data set.

knee bending vs. frontal elevation of arms), the intensity of the actions (e.g. sitting and relaxing vs. cycling), and the pace or dynamisms of their execution.

Figure 7 shows the raw data, filtered data, activity and actual activity changes. It's crucial to pick the right validation method for a given dataset. The best tactic is the hold-out validation method.

In this case, 70% of the dataset is taken into account for training, and 30% of it is used for testing. To achieve decent results, we used a hold-out validation technique [33]. Additionally, we calculated the inferred confusion Matrix for evaluating accuracy. Figure 8 shows the confusion matrix for the proposed model HARNet. The classification accuracy for the MHEALTH dataset is enhanced utilizing the suggested optimization framework, and the accuracy is increased using HARNet Architecture from 99.5% to 99.8%. Due to the accuracy of an algorithm, the HARNet deep learning framework is recommended for HAR system. As a result, the suggested prognostic tool can be used to recognize the Human activity.

4.2. Training and validation accuracy

The accuracy of the DL model can be graphically displayed using the Receiver Operating Characteristic curve. Additionally, it is used to support the validity of the proposed model. The following Figure 9 displays the ROC curve representation. The ROC curve represents the relationship between the false positive rate (FPR) and the true positive rate (TPR) for various threshold values. A typical statistic for evaluating the effectiveness of various classifiers is the area under the ROC curve (AUC). A HARNet classifier has an AUC of 1.0, meaning that higher AUC indicates better performance [34].

The effectiveness of the recommended model is compared to current techniques for HAR systems in this section. Recently, various new methods to detect human behaviour activities have been proposed by researchers. With the same starting data set, all alternative DL and ML models are compared to the proposed HARNet. The suggested model's comparison

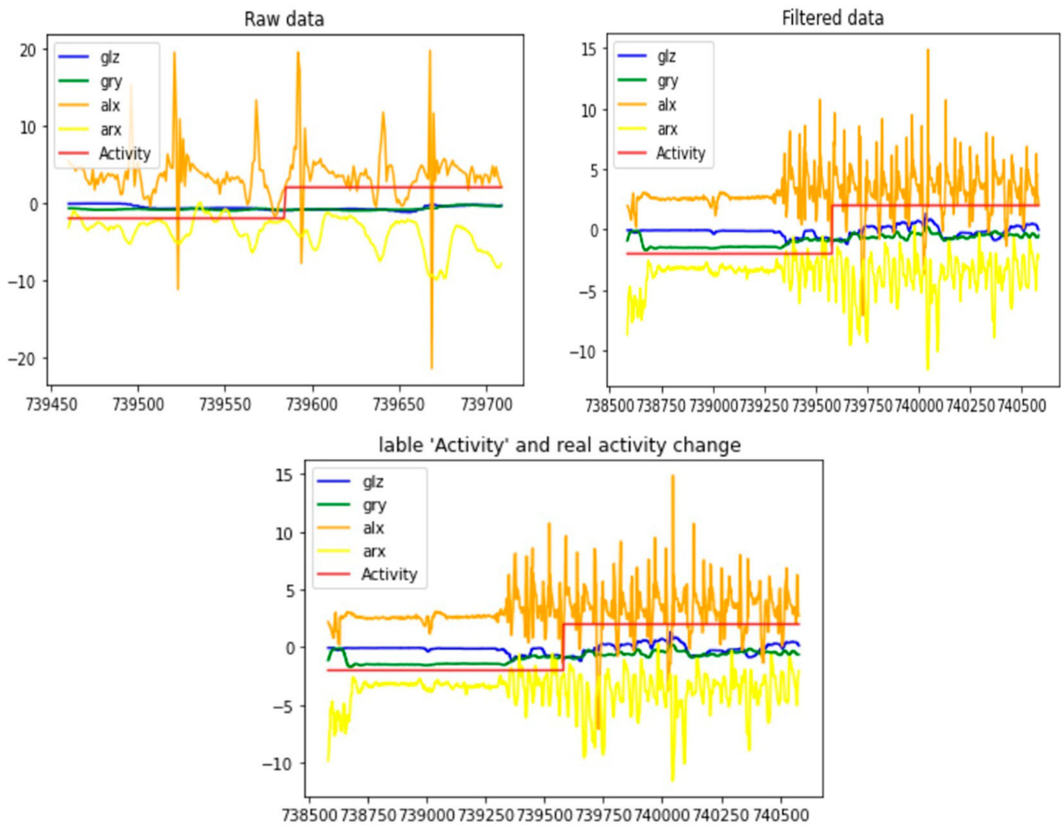

Figure 7. Visualization of raw data, filtered data and their activity changes.

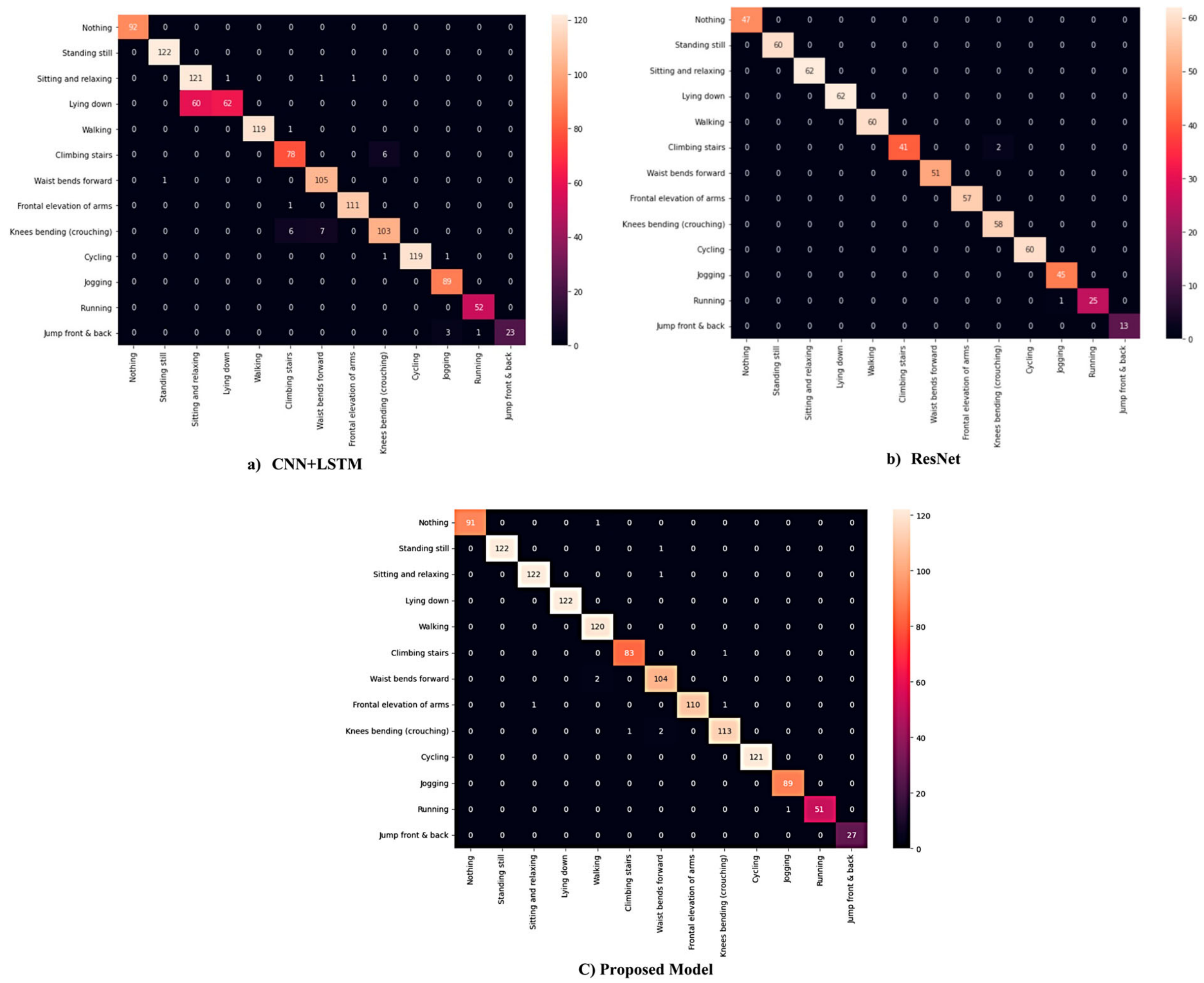

Figure 8. Confusion Matrix for CNN + LSTM, ResNet and Proposed Model.

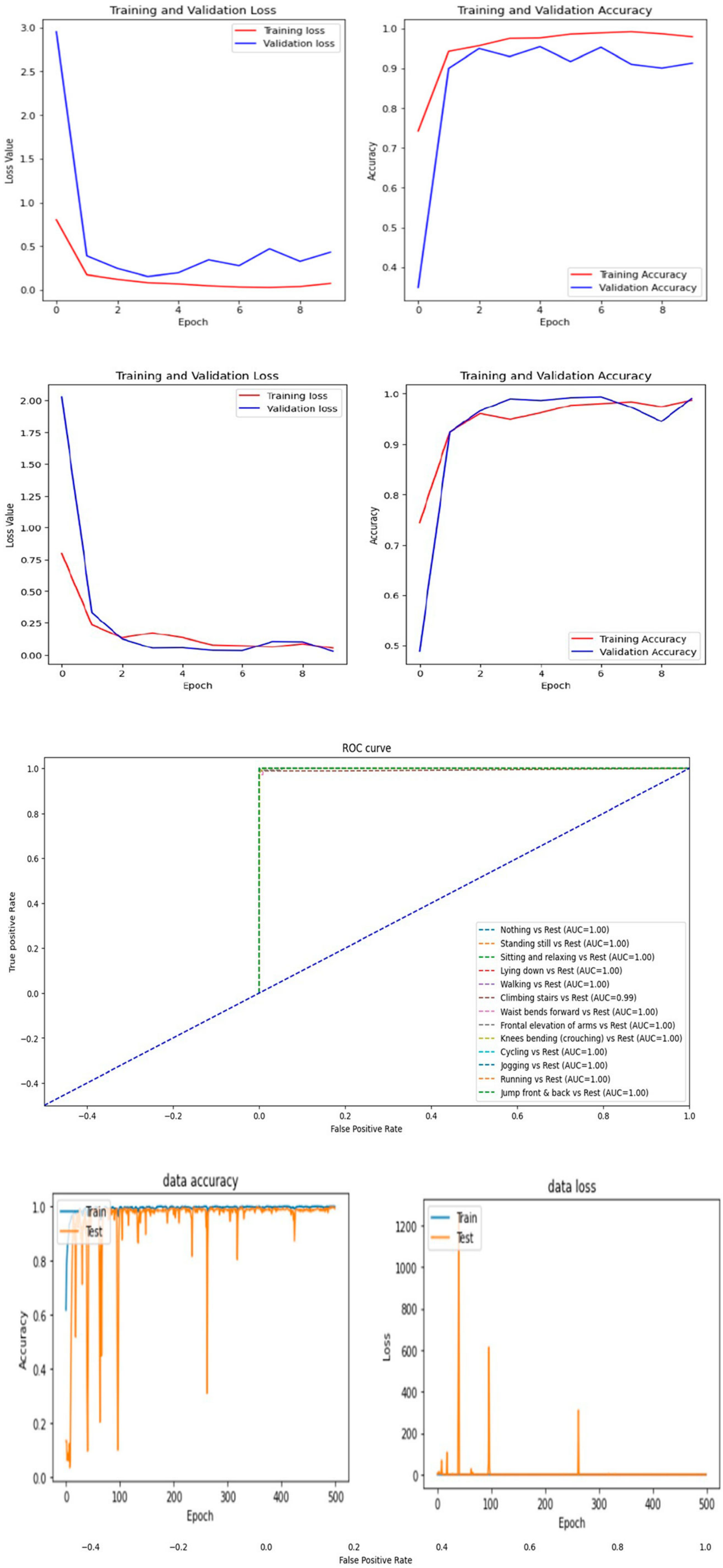

Figure 9. ROC curve for CNN + LSTM, CNN, ResNet and Proposed Model.

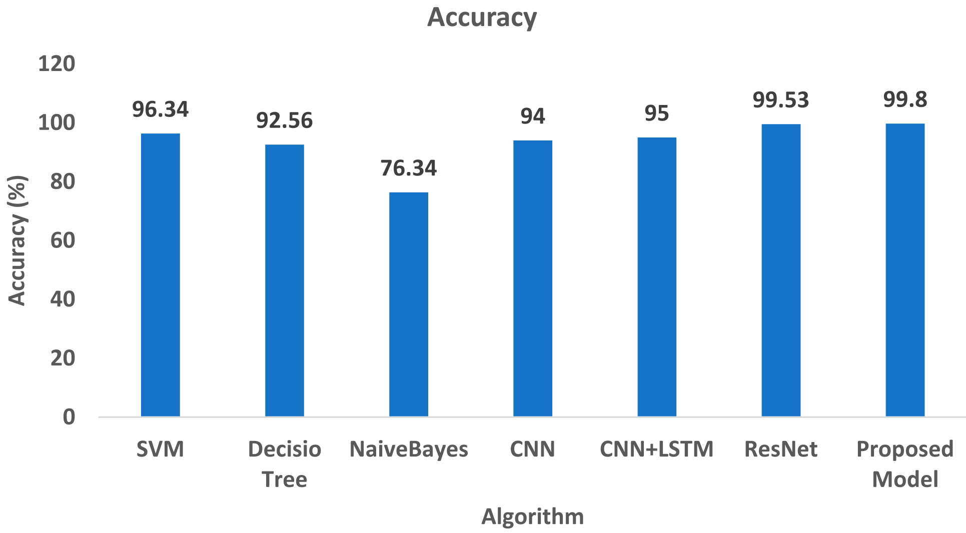

Figure 10. Comparative analysis for HARNet with other state of art algorithms.

Table 1. Summary of classification report.

Name of activity	Precision	Recall	F1-score	Support
0	1.00	0.99	0.99	92
1	1.00	0.99	1.00	123
2	0.99	0.99	0.99	123
3	1.00	1.00	1.00	122
4	0.98	1.00	0.99	120
5	0.99	0.99	0.99	84
6	0.96	0.98	0.97	106
7	1.00	0.98	0.99	112
8	0.98	0.97	0.98	116
9	1.00	1.00	1.00	121
10	0.99	1.00	0.99	89
11	1.00	0.98	0.99	52
12	1.00	1.00	1.00	27
Macro Average	0.99	0.99	0.99	1287
Weighted Average	0.99	0.99	0.99	1287

results based on performance measures like accuracy are shown in Figure 10 and the visualization of classification report for the proposed HARNet architecture is described in Table 1.

5. Conclusion

The main goal of this architecture is to create a reliable HAR system using data from sensors. In our investigation, HARNet outperformed all other machine learning algorithms, as was the case in many previous machine learning initiatives. The robust characteristics are additionally obtained for activity training using CNN and LSTM. Transfer learning is used via SVM in the LSTM's final layer. When compared to traditional ML and DL algorithms, the proposed HARNet showed significant advantages. The system's total accuracy rating is 99.8%, and it has been tested for 12 different physical activities. Future work will focus on acquiring and refining additional trustworthy traits for more accurate and complex activity recognition in real-time scenarios [24].

Disclosure statement

No potential conflict of interest was reported by the author(s).

References

[1] Jobanputra C, Bavishi J, Doshi N. Human activity recognition: a survey. Proc Comput Sci. 2019;155:698–703.

[2] Cao L, Wang Y, Zhang B, et al. GCHAR: an efficient group-based context – aware human activity recognition on smartphone. J Parallel Distrib Comput. 2018;118:67–80. doi:10.1016/j.jpdc.2017.05.007

[3] Anikwe CV, Nweke HF, Ikegwu AC, et al. Mobile and wearable sensors for data-driven health monitoring system: state-of-the-art and future prospect. Expert Syst Appl. 2022;202:117362, doi:10.1016/j.eswa.2022.117362

[4] Meegahapola L, Droz W, Kun P, et al. Generalization and personalization of mobile sensing-based mood inference models. Proc ACM Interact Mob Wear Ubiq Technol. 2022;6(4):1–32. doi:10.1145/3569483

[5] Peng W, Kanthawala S, Yuan S, et al. A qualitative study of user perceptions of mobile health apps. BMC Public Health. 2016;16(1):1–11. doi:10.1186/s12889-016-3808-0

[6] Kraft R, Birk F, Reichert M, et al. Efficient processing of geospatial mhealth data using a scalable crowdsensing platform. Sensors. 2020;20(12):3456, doi:10.3390/s20123456

[7] Hilty DM, Chan S. Human behavior with mobile health: smartphone/devices, apps and cognition. Psychol Cogn Sci. 2018;4(2):36–47. doi:10.17140/PCSOJ-4-141

[8] Miah J, Mamun M, Rahman MM, et al. Mhfit: mobile health data for predicting athletics fitness using machine learning models. In: Proceedings of the international seminar on machine learning, optimization, and data science. 2022.

[9] Liu JC, Goetz J, Sen S, et al. Learning from others without sacrificing privacy: simulation comparing centralized and federated machine learning on mobile health data. JMIR Mhealth Uhealth. 2021;9(3):e23728. doi:10.2196/23728

[10] Sannino G, Bouguila N, De Pietro G, et al. Artificial intelligence for mobile health data analysis and processing. Mob Inf Syst. 2019;2019:1–2. doi:10.1155/2019/2673463

[11] O'Halloran J, Curry E. A comparison of deep learning models in human activity recognition and behavioural prediction on the MHEALTH dataset. In: AICS; 2019. p. 212–223.
[12] Chamberlain RS, Sond J, Mahendraraj K, et al. Determining 30-day readmission risk for heart failure patients: the readmission after heart failure scale. Int J Gen Med. 2018;11:127–141. doi: 10.2147/IJGM.S150676
[13] Mo L, Li F, Zhu Y, et al. Human physical activity recognition based on computer vision with deep learning model. In: 2016 IEEE international instrumentation and measurement technology conference proceedings. IEEE; 2016. p. 1–6.
[14] Islam MM, Nipun SAA, Islam M, et al. An empirical study to predict myocardial infarction using K-means and hierarchical clustering. In: Machine learning, image processing, network security and data sciences: second international conference, MIND 2020, Silchar, India, July 30-31, 2020, Proceedings, part II 2. Singapore: Springer; 2020. p. 120–130.
[15] Al-Ghannam R, Al-Dossari H. Prayer activity monitoring and recognition using acceleration features with mobile phone. Arab J Sci Eng. 2016;41:4967–4979. doi:10.1007/s13369-016-2158-7
[16] Robert S H Istepanian. M-health 2.0 – The future of mobile health from 5G and IOT perspectives 21-22 April In 2016 2nd International Conference and business expo in wireless and telecommunications; 2016, 1–3. doi:10.4172/2090-4886.C1.006
[17] Kerdjidj O, Ramzan N, Ghanem K, et al. Fall detection and human activity classification using wearable sensors and compressed sensing. J Ambient Intell Humaniz Comput. 2020;11:349–361. doi:10.1007/s12652-019-01214-4
[18] Maszczyk A. A neural network model approach to athlete selection. Sports Eng; 13:83–93. doi:10.1007/s12283-010-0055-y
[19] Dahou A, Al-qaness MA, Abd Elaziz M, et al. Human activity recognition in IoHT applications using arithmetic optimization algorithm and deep learning. Measurement (Mahwah N J). 2022;199:111445, doi:10.1016/j.measurement.2022.111445
[20] Osco LP, Junior JM, Ramos APM, et al. (2021). A review on deep learning in UAV remote sensing. arXiv preprint arXiv:2101 10861.
[21] Banos O, Garcia R, Holgado-Terriza JA, et al. Mhealthdroid: a novel framework for agile development of mobile health applications. In: Ambient assisted living and daily activities: 6th international work-conference, IWAAL 2014, Belfast, UK, December 2–5, 2014. proceedings: Springer International; 2014; Vol. 6 pp. 91–98.
[22] Banos O, Villalonga C, Garcia R, et al. Design, implementation and validation of a novel open framework for agile development of mobile health applications. Biomed Eng Online. 2015;14(2):1–20.
[23] Saxena A. An introduction to convolutional neural networks. Int J Res Appl Sci Eng Technol. 2022;10(12):943–947. doi:10.22214/ijraset.2022.47789
[24] Ajit A, Acharya K, Samanta A. A review of convolutional neural networks. In: 2020 international conference on emerging trends in information technology and engineering (ic-ETITE). New York: IEEE; 2020.
[25] Chaabane SB, Hijji M, Harrabi R, et al. Face recognition based on statistical features and SVM classifier. Multimed Tools Appl. 2022;81(6):8767–8784. doi:10.1007/s11042-021-11816-w
[26] Kiruthika NS, Thailambal DG. Dynamic light weight recommendation system for social networking analysis using a hybrid LSTM-SVM classifier algorithm. Optic Memory Neural Netw. 2022;31(1):59–75. doi:10.3103/S1060992X2201009X
[27] Ariza-Colpas PP, Vicario E, Oviedo-Carrascal AI, et al. Human activity recognition data analysis: history, evolutions, and new trends. Sensors. 2022;22(9):3401, doi:10.3390/s22093401
[28] Patle A, Chouhan DS. SVM kernel functions for classification. In: 2013 International Conference on Advances in Technology and Engineering (ICATE). IEEE; 2013. p. 1–9.
[29] Dahou A, Al-qaness MA, Abd Elaziz M, et al. Human activity recognition in IoHT applications using arithmetic optimization algorithm and deep learning. Measurement (Mahwah N J). 2022;199:111445, doi:10.1016/j.measurement.2022.111445
[30] Ramanujam E, Perumal T, Padmavathi S. Human activity recognition with smartphone and wearable sensors using deep learning techniques: a review. IEEE Sens J. 2021;21(12):13029–13040. doi:10.1109/JSEN.2021.3069927
[31] Zhang Z. Improved adam optimizer for deep neural networks. In: 2018 IEEE/ACM 26th international symposium on quality of service (IWQoS). IEEE; 2018. p. 1–2.
[32] Priyadarshini I, Sharma R, Bhatt D, et al. Human activity recognition in cyber-physical systems using optimized machine learning techniques. Cluster Comput. 2023;26(4):2199–2215. doi:10.1007/s10586-022-03662-8
[33] Wu J, Liang H, Ding C, et al. Improving the accuracy in classification of blood pressure from photoplethysmography using continuous wavelet transform and deep learning. Int J Hypertens. 2021;volume 2021:17–19. doi:10.1155/2021/9938584
[34] Powers DM. Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation. arXiv Preprint ArXiv. 2020: 2010.16061.