J. Electrical Systems 20-7s (2024): 890-901

[1] Aniket Bandyopadhyay
[2] Manoj Kumar

# Gait-Based Human Activity Recognition Using Efficient Sensor Fusion and Deep Learning Approach

Journal of Electrical Systems

***Abstract: -*** Human activity recognition is an important area of computer vision research. Its applications include surveillance systems, patient monitoring systems, and a variety of systems that involve interactions between persons and electronic devices such as human-computer interfaces. Most of these applications require an automated recognition of high-level activities, composed of multiple simple (or atomic) actions of persons. A novel feature selection approach is then proposed in order to select a subset of discriminant features, construct an online activity recognizer with better generalization ability, and reduce the smartphone power consumption. Experimental results on a publicly available data set show that the fusion of both accelerometer and gyroscope data contributes to obtain better recognition performance than that of using single source data, and that the proposed feature selector outperforms three other comparative approaches in terms of four performance measures. Such activity profiling systems are dependent on classification algorithms which can effectively interpret body-worn sensor data and identify different activities. the key research challenges that human activity recognition shares with general pattern recognition and identify those challenges that are specific to human activity recognition. The aim of paper is to explore real life applications like contactless employee recognition system using gait analysis which uses sensor data as base to identify employees based on their gait movement. This requires understanding the dimensions of sensor data and its application exploring other potential real-life applications and optimizing the methodology are also one of the core objectives.

***Keywords:*** Machine Learning, Deep Learning, CNN, CNN-LSTM, Human Activity Recognition, UCI.

## I. INTRODUCTION

Human activity recognition (HAR) has multifaceted applications due to its worldly usage of acquisition devices such as smartphones, video cameras, and its ability to capture human activity data. While electronic devices and their applications are steadily growing, the advances in Artificial intelligence (AI) have revolutionized the ability to extract deep hidden information for accurate detection and its interpretation. This yields a better understanding of rapidly growing acquisition devices, AI, and applications, the three pillars of HAR under one roof. There are many review articles published on the general characteristics of HAR, a few have compared all the HAR devices at the same time, and few have explored the impact of evolving AI architecture. In our proposed review, a detailed narration on the three pillars of HAR is presented covering the period from the $19^{th}$ century to present day.

Further, the review presents the recommendations for an improved HAR design, its reliability, and stability. Five major findings were: (1) HAR constitutes three major pillars such as devices, AI and applications; (2) HAR has dominated the healthcare industry; (3) Hybrid AI models are in their infancy stage and needs considerable work for providing the stable and reliable design. Further, these trained models need solid prediction, high accuracy, generalization, and finally, meeting the objectives of the applications without bias; (4) little work was observed in abnormality detection during actions; and (5) almost no work has been done in forecasting actions. We conclude that: (a) HAR industry will evolve in terms of the three pillars of electronic devices, applications, and the type of AI. (b) AI will provide a powerful impetus to the HAR industry in future. One of our paper's aims is to detect Employees from the company based on their gait analysis and use the Machine Learning Algorithm to detect which of the employee it is and based on it open the gates of the Company building automatically if the employee belongs to the authorized database and matches their gait data [1-5].

Gait analysis is the systematic study of human walking patterns or the way people move when they walk. Observational gait analysis is clinically useful with videotape slow-motion replay and freeze-frame, offering significant improvement over unaided visual observation. Any form of observational gait analysis, however, has limited precision and is more descriptive than quantitative [6]. It is an important field of study in various disciplines, including biomechanics, physical therapy, sports science, and orthopedics. Gait analysis involves the collection and

[1, 2] School of Computer Science and Engineering VIT Bhopal University Bhopal-Indore Highway, Kothrikalan, Sehore, Madhya Pradesh 466114.

Email: [1] aniket.bandyopadhyay2019@vitbhopal.ac.in, [2] manojkumar@vitbhopal.ac.in

interpretation of data related to a person's walking or running stride to better understand their movement patterns, detect abnormalities, and make informed decisions for diagnosis, treatment, or performance improvement. Walking also requires motor control and motor coordination. Many orthopedic surgical procedures are designed to improve ambulation by optimizing joint forces, thereby alleviating, or preventing pain and improving energy conservation. Gait analysis, accomplished by either simple observation or three-dimensional analysis with measurement of joint angles (kinematics), joint forces (kinetics), muscular activity, foot pressure, and energetics (measurement of energy utilized during an activity), allows the physician to design procedures tailored to the individual needs of patients [7]. This shows its efficacy in the field of medicine and its contribution in it.

Gait analysis has expanding indications in orthopedic surgery, both for clinical and research applications. Early work has been particularly helpful for understanding pathologic gait deviations in neuromuscular disorders and biomechanical imbalances that contribute to injury [8]. As such understanding gait analysis could lead to predictive analysis.

### *1.1 key aspects of gait analysis:*

**Data Collection**: Gait analysis typically involves collecting data from various sources, including video recordings, force platforms, electromyography (EMG), and other sensors. High-speed cameras and motion-capture systems are often used to track the movement of specific body markers.

**Parameters Measured**: Various parameters are measured during gait analysis, including step length, step width, stride length, walking speed, joint angles, joint forces, ground reaction forces, and temporal patterns. These measurements help in understanding the mechanics of walking.

**Clinical Use**: There is still not an accepted general theory of why we walk the way we do. In the absence of this, many explanations of walking address the mechanisms by which specific movements are achieved by particular muscles. A whole new methodology is developing to determine the functions of individual muscles [9]. This should only progress over the years and help us in getting a better understanding of its application. Below are some Criteria as used in Clinical data collection:

| |
|---|
| Reproducible |
| Stable (independent of mood, motivation and pain) |
| Accurate |
| Appropriately validated |
| Capable of distinguishing between normal and abnormal |
| Must not alter the function it is measuring |
| Reported in form analogous to accepted clinical concepts |
| Cost-effective |
| Not observable by the skilled clinician |

**Figure 1.** Criteria as mentioned [9]

**Concern Regarding Gait Data**: The data produced from gait analysis, however, is not necessarily free of errors. (i) to estimate the errors associated with quantitative gait data; and (ii) to propose a method for incorporating the knowledge of these errors into the clinical interpretation process [10]. We consider these errors and circumvent it with more data and cross reference. This gait representation is based on simple features such as moments extracted from orthogonal view video silhouettes of human walking motion [11]. Those gathered from sensor data. There are three most easily measured general gait parameters: (1) cadence, (2) velocity, and (3) stride length [12].

**Human Activity Recognition**: Recognizing human activities from video sequences or still images are a challenging task due to problems, such as background clutter, partial occlusion, changes in scale, viewpoint, lighting, and appearance. Many applications, including video surveillance systems, human-computer interaction, and robotics for human behavior characterization, require a multiple activity recognition system [13]. Here we have considered six actions to identify. Machine learning techniques like decision trees, K-nearest neighbors, support vector machines, hidden Markov models are reviewed for HAR and later the survey for deep neural network techniques like artificial neural networks, convolutional neural networks and recurrent neural networks is also presented [14].

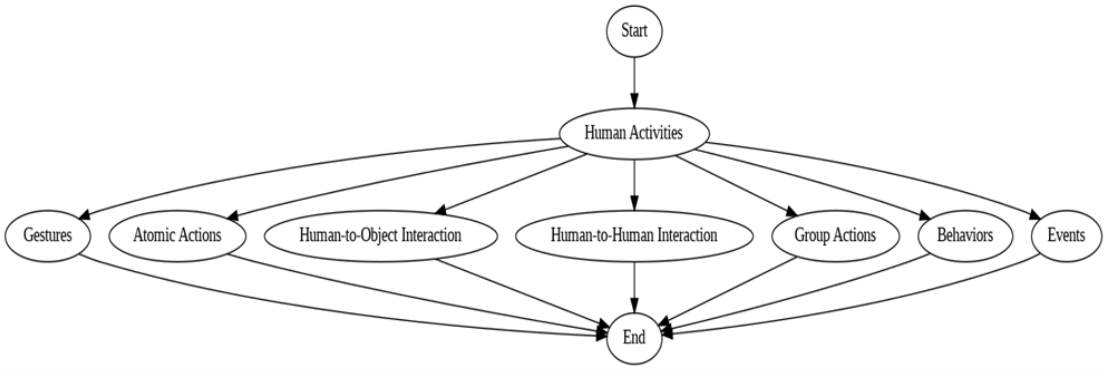



**Figure 2**

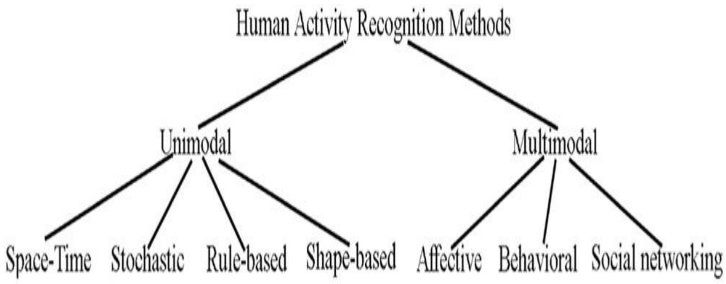



**Figure 3** [13]

There has been research focused on CCTV videos and camera images to detect human poses using HAAR Feature-based Classifier and recognize the activities of the human using the Convolutional Neural Network (CNN) Classifier. One of which had a Human Activity Recognition System which was trained using own collected dataset which is composed of 5648 images. The approach accomplished an efficacious detection accuracy of 99.86% and recognition accuracy of 99.82% with approximately 22 frames/second after 20 epochs [15]. In many recent works, the recognition model architecture uses CNN and long short-term memory units (LSTM) - attention models to extract spatial and temporal features from the input video [16]. This is one of our main models which we will showcase as well as Support Vector Machine (SVM) to identify individual static activity and 1D Convolutional Neural Network (CNN)-based deep learning model for individual moving activity recognition [17] is also used in machine learning prediction. Some papers have used a Convolutional Neural Network (CNN) and using the 2D pose estimation technique to the system [18].

The MPJA-HAD dataset also provides joint angle changes from each of 15 body positions. Joint angles directly relate to human activities performing and experimental results show the competitiveness of the created dataset in HAR tasks [19]. The architectures of smart devices will need to interpret human action in real-time and predict humans' immediate intention in complex, noisy and cluttered environments [20]. This is one of our objectives of work. Despite human activity recognition (HAR) being an active field for more than a decade, there are still key aspects that, if addressed, would constitute a significant turn in the way people interact with mobile devices [21]. Using this device, we collect our sensor gait data for predicting the action done by the person.

**Structure of paper:** Section one contains introduction about activity recognition. Literature review related to work has been done in section 2. Section three contains proposed methodology and dataset used. Results and discussion have been done in section four. Paper conclusion and future scope is the part of section five**.**

## II. RELATED WORK:

**Table 1.** Comparative Study of Related Work Done in Activity Recognition **[1-31]**

| Author & Year | Published in | Concept discussed | Limitations |
|---|---|---|---|
| Aggarwak,J.K. (2011) [1] | ACM Computing surveys | Human activity analysis | Application of data |

| | | | |
|---|---|---|---|
| Chen, Y., Xue,Y , (2012) [2] | Distributed Sensor networks | Human activity recognition | Sensor data acquisition |
| Preece, S.J., Goulermas (2009) [3] | Physiological Measurement | Activity recognition | Body-Mounted Sensors accuracy |
| Bulling, A., Blanke (2014) [4] | ACM Computing surveys | Activity recognition | Inertial Sensors Tutorial usability |
| Khan, Mau, Gadekallu, T.R. (2020) [5] | Sensors | Human activity recognition | Wearable Sensors Survey |
| GF Harris, JJ Wertsch (1994) [6] | Elsevier | Gait Behavior | Gait Analysis Methodology |
| Chambers, Henry G. MD (2002) [7] | Journal of the American Academy of Orthopaedic Surgeons | Gait Analysis | Practical Gait Guide |
| Hecht, Garin G. MD (2002) [8] | Journal of the American Academy of Orthopaedic Surgeons | Gait Analysis | Gait Analysis History |
| Baker, R. (2006) [9] | NeuroEngineering Rehabil | Gait Analysis methods | Rehabilitation Gait Methods |
| MH Schwartz, JP Trost (2004) [10] | Elsevier | Gait and posture | Quantitative Gait Errors |
| L. Lee and W. E. L. Grimson (2002) [11] | Proceedings of Fifth IEEE International Conference on Automatic Face Gesture Recognition | Gait Analysis | Gait Recognition Of Actions based on data and Accuracy |
| MW Whittle (1993) [12] | Elsevier | Gait Analysis | Gait Analysis Soft Tissues |
| M Vrigkas, C Nikou (2015) [13] | Frontiers in Robotics and AI | Human Activity recognition | Activity Recognition Methods |
| C Jobanputra, J Bavishi (2019) [14] | Elsevier | Procedia computer Science | Human Activity Survey |
| L.Xie, J. Tian (2018) [15] | INERTIAL | Human activity recognition | Inertial Sensor Barometer |
| N. Archana, K. Hareesh (2021) [16] | ACCESS | Real-time human activity recognition | Accuracy of Output |
| M. M. Hossain Shuvo, N. Ahmed (2020) [17] | AIPR | Hybrid approach for human activity recognition | SVM CNN Recognition Efficiency |
| A. S. Dillip, N. S. S. (2022) [18] | WiSPNET | Suspicious Human activity recognition | Pose CNN Recognition Efficiency |
| O. D. Lara and M. A. Labrador (2013) [19] | Communication Surveys & Tutorials | Human activity recognition using sensors | Wearable Sensors Survey |
| P. Nikolov, O. Boumbarov (2018) [20] | TSP | Skelton-based Human activity recognition | Skeleton-Based Recognition |
| H.Yang, X.Wen (2022) [21] | ICAMechs | Human Activity Recognition Using Wearable Sensors | Joint Angle Dataset |

| | | | |
|---|---|---|---|
| Anguita, D., Ghio (2013) [22] | Computational Intelligence and Machine learning | Human Activity Recognition using Smartphones | Smartphone Dataset |
| Reyes-Ortiz, J. L. (2016) [23] | Neurocomputing | human activity recognition using smartphones | Smartphone Transition Recognition |
| Ronao, C. A., & Cho, S. B. (2016) [24] | Expert Systems with Applications | Human activity recognition with smartphone sensors | Sensor Data Acquisition |
| Ronao, C. A., & Cho, S. B. [(2016) [25] | Neural Networks | Deep convolutional neural networks for human activity recognition | DL Human Activity Recognition |
| Reyes-Ortiz, J. L. (2016) [26] | Neurocomputing | Transition-aware human activity recognition using smartphones | Transition-Aware Recognition Efficacy |
| S. Hochreiter and J. Schmidhuber (1997) [27] | Neural Computation | "Long Short-Term Memory," | Early LSTM without CNN |
| T. J. Chin, L. Wang, K. Schindler (2006) [28] | Int. Conf. Image Process | Extrapolating learned manifolds for human activity recognition | Reduced Accuracy |
| Y. L. Hsu, S. L. Lin, P. H. Chou (2017) [29] | ICASI 2017 | Application of nonparametric weighted feature extraction for an inertial-signal-based human activity recognition system | Accuracy of Output in Application |
| J. YANG, J. CHENG, and H. LU (2009) [30] | Icme 2009 | Human Activity Recognition Based On The Blob Features National Laboratory of Pattern Recognition | Accuracy of Results |
| M. . Arshad (2021) [31] | EAI Endorsed Trans IoT | Hybrid Machine Learning Techniques to detect Real Time Human Activity using UCI Dataset | Accuracy Of Data |

A literature review on human activity recognition (HAR) and gait analysis provides insights into the advancements, challenges, and applications of these fields. Below is a comprehensive overview of key research papers and trends in both areas:

**Human Activity Recognition (HAR):**

**1. Survey Papers:**

- Anguita, D., Ghio, A., Oneto, L., Parra, X., & Reyes-Ortiz, J. L. (2013). A Public Domain Dataset for Human Activity Recognition Using Smartphones. 21st European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning[22].
- Reyes-Ortiz, J. L., Oneto, L., Samà, A., Parra, X., & Anguita, D. (2016). Transition-aware human activity recognition using smartphones. Neurocomputing, 171, 754-767[23].

**2. Deep Learning Approaches:**

- Ronao, C. A., & Cho, S. B. (2016). Human activity recognition with smartphone sensors using deep learning neural networks. Expert Systems with Applications, 59, 235-244[24].
- Ronao, C. A., & Cho, S. B. (2016). Deep convolutional neural networks for human activity recognition with smartphone sensors. Neural Networks, 78, 137-147[25].

**3. Sensor Fusion Techniques:**

- Reyes-Ortiz, J. L., Oneto, L., Samà, A., Parra, X., & Anguita, D. (2016). Transition-aware human activity recognition using smartphones. Neurocomputing, 171, 754-767[26].
- Ronao, C. A., & Cho, S. B. (2016). Human activity recognition with smartphone sensors using deep learning neural networks. Expert Systems with Applications, 59, 235-244.


**Objective of work:**

To Identify a contactless employee recognition system which can Identify employees of its company using sensor data by analysing their gait behaviour and associate that gait behaviour with a person.
The UCI HAR dataset is a popular dataset used in the field of machine learning and human activity recognition, which contains sensor data collected from smartphones during various physical activities. Here's an overview of some key points and findings from relevant literature:

**Dataset Description:**

The UCI HAR dataset contains data from the accelerometers and gyroscope of smartphones worn by participants performing six different activities: walking, walking upstairs, walking downstairs, sitting, standing, and laying. Each data point is labeled with the corresponding activity.

**Data Preprocessing:**

Many papers discuss the importance of preprocessing the dataset, including data cleaning, feature selection, and normalization. Techniques like feature scaling and dimensionality reduction are commonly applied.

**Feature Engineering:**

Researchers often extract a variety of features from the raw sensor data. These features might include mean, standard deviation, time-domain features, frequency-domain features, and more. Feature selection and engineering play a crucial role in model performance.

**Classification Algorithms:**

Researchers commonly apply various machine learning and deep learning algorithms to classify the activities. These algorithms include Decision Trees, Random Forests, Support Vector Machines, K-Nearest Neighbors, and neural networks, such as Convolutional Neural Networks (CNNs) or Long Short-Term Memory (LSTM) networks.

**Performance Evaluation:**

Evaluation metrics used in the literature typically include accuracy, precision, recall, F1-score, and confusion matrices. Cross-validation and training-test splits are often used to assess model performance.

**Deep Learning Approaches:**

Some studies focus on the application of deep learning techniques to the UCI HAR dataset. Convolutional neural networks (CNNs) and recurrent neural networks (RNNs) have been used to achieve state-of-the-art results in human activity recognition and our aim is to implement these approaches with the appropriate methodology.

**Ensemble Methods:**

Ensemble methods, such as Random Forest and Gradient Boosting, are employed to combine multiple classifiers for improved accuracy and robustness. These can be used to gain an insight of the type and its properties of the data and how it works with the methods.

**Transfer Learning:**

Some researchers explore the use of transfer learning, where models pretrained on large datasets are fine-tuned for human activity recognition using the UCI HAR dataset. This can help in cases with limited labeled data.

**Real-World Applications:**

Beyond academic research, the UCI HAR dataset has found applications in real-world scenarios, such as healthcare, fitness tracking, and security, where activity recognition is essential.

**Challenges and Future Directions:**

Many papers discuss challenges in human activity recognition, such as handling noisy sensor data, dealing with class imbalances, and improving model interpretability. Future directions often involve investigating more complex activities, larger datasets, and more robust models.
Overall, the UCI HAR dataset has served as a valuable benchmark for developing and testing human activity recognition models. Research papers on this dataset have contributed to the advancement of techniques in this field and have practical implications in areas like healthcare, sports, and assistive technologies.

- **Research Gap and Problem Identification**

To build a contactless employee security system based on the employees' historical gait analysis.

This requires two crucial steps.
1. To identify the gait activity based on the data from gyroscope and accelerometer.
2. Matching the identified gait data with historical gait data of the employee to recognize them.

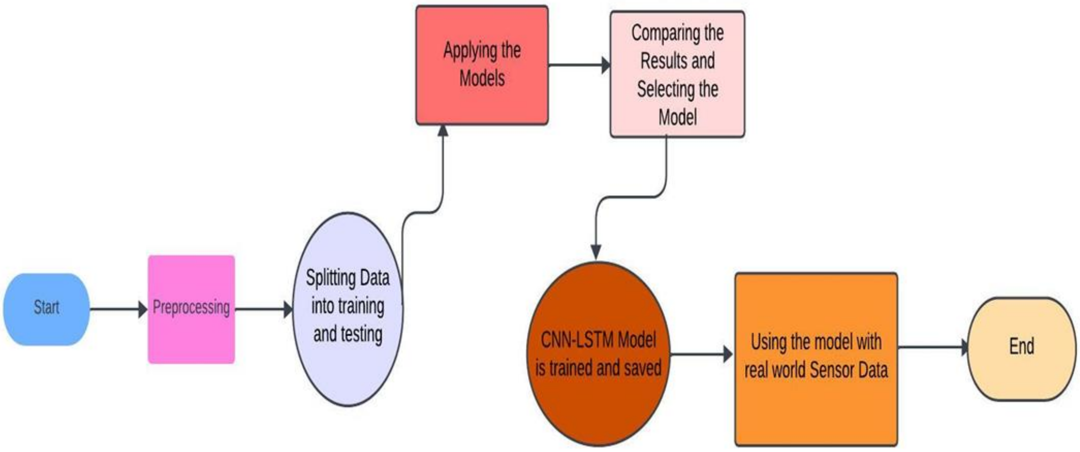



**Figure 4.** (Proposed Framework)

**Research Gap:**
1. Limited Evaluation in Real-World Settings: Many studies using the UCI HAR dataset focus on controlled environments. A research gap may exist in assessing the generalizability and performance of models in real-world, uncontrolled settings where sensor data can be noisier.
2. Transfer Learning: There may be room for research into transfer learning techniques that leverage models trained on larger and more diverse datasets to improve the recognition of activities in the UCI HAR dataset, particularly for uncommon or complex activities.
3. Exploring Alternative Sensor Data: The UCI HAR dataset primarily utilizes accelerometer and gyroscope data from smartphones. Research could explore the integration of other sensors (e.g., GPS, magnetometer) or additional data sources (e.g., audio or image data) to enhance activity recognition accuracy.

## III. Proposed Methodology and Dataset

### *3.1 Dataset description:*

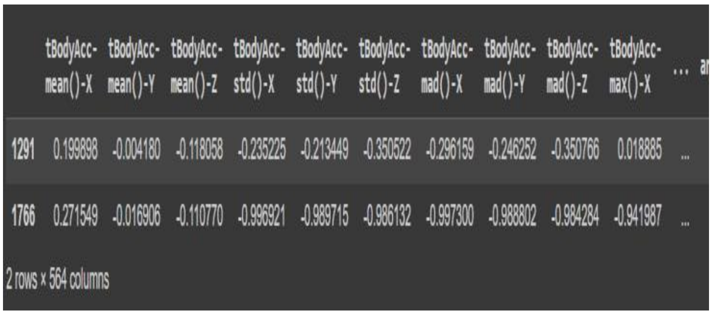
This is a Pre-processed dataset. The pre-processing steps included: Pre-processing accelerometer and gyroscope using noise filters. Splitting data into fixed windows of 2.56 seconds (128 data points) with 50% overlap. Splitting of accelerometer data into gravitational (total) and body motion components. Feature engineering was applied to the window data, and a copy of the data with these engineered features was made available. A number of time and frequency features commonly used in the field of human activity recognition were extracted from each window. The result was a 561-element vector of features.
Train Data:

| | tBodyAcc-mean()-X | tBodyAcc-mean()-Y | tBodyAcc-mean()-Z | tBodyAcc-std()-X | tBodyAcc-std()-Y | tBodyAcc-std()-Z | tBodyAcc-mad()-X | tBodyAcc-mad()-Y | tBodyAcc-mad()-Z | tBodyAcc-max()-X | ... |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 6878 | 0.277417 | -0.016868 | -0.108601 | -0.993599 | -0.996214 | -0.997417 | -0.993911 | -0.99595 | -0.997326 | -0.937886 | ... |
| 3442 | 0.275638 | -0.015203 | -0.122476 | -0.997891 | -0.991359 | -0.989268 | -0.997980 | -0.99071 | -0.987728 | -0.945770 | ... |

2 rows × 564 columns

**(a)**

Test Data:



| | tBodyAcc-mean()-X | tBodyAcc-mean()-Y | tBodyAcc-mean()-Z | tBodyAcc-std()-X | tBodyAcc-std()-Y | tBodyAcc-std()-Z | tBodyAcc-mad()-X | tBodyAcc-mad()-Y | tBodyAcc-mad()-Z | tBodyAcc-max()-X | ... |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1291 | 0.199898 | -0.004180 | -0.118058 | -0.235225 | -0.213449 | -0.350522 | -0.296159 | -0.246252 | -0.350766 | 0.018885 | ... |
| 1766 | 0.271549 | -0.016906 | -0.110770 | -0.996921 | -0.989715 | -0.986132 | -0.997300 | -0.988802 | -0.984284 | -0.941987 | ... |

2 rows × 564 columns

**(b)**

**Figure 5.** Train and Test data sample

### *3.2 Proposed Framework:*

The proposed framework showcases an overview of the steps taken splitting the data into training and testing to prepare the data for evaluating models and results of the data, gathering conclusions from this, we move forward with the CNN-LSTM model and then use this model with real world sensor data in real time.

A Multivariate Time Series consist of more than one-time dependent variable and each variable depends not only on its past values but also has some dependency on other variables.

To deal with MTS, one of the most popular methods is Vector Auto Regressive Moving Average models (VARMA) that is a vector form of autoregressive integrated moving average (ARIMA) that can be used to examine the relationships among several variables in multivariate time series analysis.

LSTM: Long-short term memory (LSTM)is a deep learning model; it is a kind of Recurrent Neural Network (RNN) to read the sequence dependencies. We have used CNN-LSTM. Which is made for reading long sequences of 1D time inputs with multiple features associated with that particular time frame.

The final step - The new Model will be trained on the real time data which we get after the PCA feature selection process. And would require the testing of machine learning models using the data to predict the action.

## IV. Results and Discussion

**Table 2.** Comparative study of related work

| Models used | Recall | Precision | F1-score | Accuracy |
|---|---|---|---|---|
| Random Forest | 1.0 | 1.0 | 1.0 | 0.922 |
| Bagging Classifier | 1.0 | 1.0 | 1.0 | 0.903 |
| AdaBoost | 1.0 | 1.0 | 1.0 | 0.531 |
| Naïve Bayes Classifier | 0.60 | 0.98 | 0.74 | 0.770 |
| Support Vector Machine | 1.0 | 1.0 | 1.0 | 0.95 |
| XGBoost | 1.0 | 1.0 | 1.0 | 0.939 |
| K-Nearest Neighbour | 0.99 | 1.0 | 0.99 | 0.900 |
| Multi-Layer Perceptron | 0.99 | 1.0 | 0.99 | 0.945 |

From this we can understand how machine learning models fare against each other in terms of accuracy it's important to note that Naïve Bayes and Adaboost don't do as well as compared to other ML models. The main aim of boosting is to decrease bias, not variance. As such it provides no benefit and having no "weaker performance" to work with does not help its accuracy. The best performing models in our studies were SVM AND MLP.

**Table 3**

| Model | Train Accuracy (%) | Test Accuracy (%) | Loss (%) | |
|---|---|---|---|---|
| LSTM | 92 | 89 | Train | Test |
| | | | 24 | 50 |
| CNN-LSTM | 96 | 91 | 10 | 76 |

**LSTM:**

```
LSTM Training Accuracy
230/230 [==============================] - 3s 13ms/step - loss: 0.2447 - accuracy: 0.9283
Accuracy:  0.9283187985420227
Loss:  0.24472583830356598
230/230 [==============================] - 3s 13ms/step
[[1224    2    0    0    0    0]
 [   0 1073    0    0    0    0]
 [   0    5  981    0    0    0]
 [   0    0    0 1206   80    0]
 [   0    0    0  257 1117    0]
 [   0    0    0  183    0 1224]]
```

**Figure 6**

STM can learn to bridge minimal time lags in excess of 1000 discrete-time steps by enforcing constant error flow through constant error carousels within special units. Multiplicative gate units learn to open and close access to the constant error flow. LSTM is local in space and time; its computational complexity per time step and weight is O. 1. Our experiments with artificial data involve local, distributed, real-valued, and noisy pattern representations. In comparisons with real-time recurrent learning, back propagation through time, recurrent cascade correlation, Elman nets, and neural sequence chunking, LSTM leads to many more successful runs, and learns much faster. LSTM also solves complex, artificial long-time-lag tasks that have never been solved by previous recurrent network algorithms [27].

The CNN LSTM architecture involves using Convolutional Neural Network (CNN) layers for feature extraction on input data combined with LSTMs to support sequence prediction. The model reads subsequences of the main sequence as blocks, extracts feature then uses the model to interpret the features. We have split each window of 128-time steps into 4 subsequences for the CNN model to process. We can then define a CNN model that expects to read in sequences with a length of 32-time steps and nine features.

The entire CNN model can be wrapped in a Time Distributed layer to allow to read in each of the four subsequences. The extracted features are then flattened and provided to the LSTM model to read, extracting its own features before a final mapping to an activity.

```
Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 time_distributed (TimeDistr  (None, None, 30, 64)     1792
 ibuted)

 time_distributed_1 (TimeDis  (None, None, 28, 64)     12352
 tributed)

 time_distributed_2 (TimeDis  (None, None, 28, 64)     0
 tributed)

 time_distributed_3 (TimeDis  (None, None, 14, 64)     0
 tributed)

 time_distributed_4 (TimeDis  (None, None, 896)        0
 tributed)

 lstm (LSTM)                 (None, 32)                118912

 dropout_1 (Dropout)         (None, 32)                0

 dense (Dense)               (None, 32)                1056

 dense_1 (Dense)             (None, 6)                 198

=================================================================
Total params: 134,310
Trainable params: 134,310
Non-trainable params: 0
_________________________________________________________________
```

**Figure 7**

**CNN-LSTM Model:**

```
CNN-LSTM Training Accuracy
230/230 [==============================] - 2s 8ms/step - loss: 0.1060 - accuracy: 0.9614
Accuracy:  0.9613710641860962
Loss:  0.10595442354679108
230/230 [==============================] - 3s 9ms/step
[[1226    0    0    0    0    0]
 [  63 1010    0    0    0    0]
 [   2    0  984    0    0    0]
 [   2    0    0 1218   66    0]
 [   0    0    0  151 1223    0]
 [   0    0    0    0    0 1407]]
```

**Figure 8**

**Table 4:** Comparative analysis of different work done with current work

| Models/ techniques used | Dataset used | Test Accuracy (%) | Ref. |
|---|---|---|---|
| Random Forests | UCI HAR | 78 | [28] |
| SVM with PCA | HAR | 85.4 | [29] |
| CNN | HAR | 82 | [30] |
| Hybrid Approach SVM-KNN | UCI HAR | 87 | [31] |
| CNN-LSTM | UCI HAR | 91 | |

```
CNN-LSTM Testing Accuracy
93/93 [==============================] - 1s 7ms/step - loss: 0.7681 - accuracy: 0.9104
Accuracy:  0.910417377948761
Loss:  0.7681483626365662
93/93 [==============================] - 0s 4ms/step
[[493   0   3   0   0   0]
 [ 14 432  25   0   0   0]
 [ 23   2 395   0   0   0]
 [  1   3   0 393  94   0]
 [  2   0   0  70 460   0]
 [  0  27   0   0   0 510]]
```

**Figure 9**

- No. of steps – 32
- Filters – 64
- Kernel size – 3
- Epochs – 30
- Batch size – 16

**Table 5:**

| | Precision | Recall | F1-score | support |
|---|---|---|---|---|
| 0 | 0.92 | 0.99 | 0.96 | 496 |
| 1 | 0.93 | 0.92 | 0.92 | 471 |
| 2 | 0.93 | 0.94 | 0.94 | 420 |
| 3 | 0.85 | 0.80 | 0.82 | 491 |
| 4 | 0.83 | 0.86 | 0.85 | 532 |
| 5 | 1.00 | 0.95 | 0.97 | 537 |

**Table 6:**

| | | | | |
|---|---|---|---|---|
| Accuracy | | | 0.91 | 2947 |
| Macro avg | 0.91 | 0.91 | 0.91 | 2947 |
| weighted | 0.91 | 0.91 | 0.91 | 2947 |

Precision – Quality of a positive prediction made by the model.
Recall – Measures the model's ability to detect Positive samples.
F1-score – combines the precision and recall scores of a model.
Support – the number of actual occurrences of the class in the specified dataset.
Training accuracy is 96%.
Testing accuracy is 91%.
There is a bit of overfitting involved in this model since the training accuracy is slightly higher than the testing accuracy.

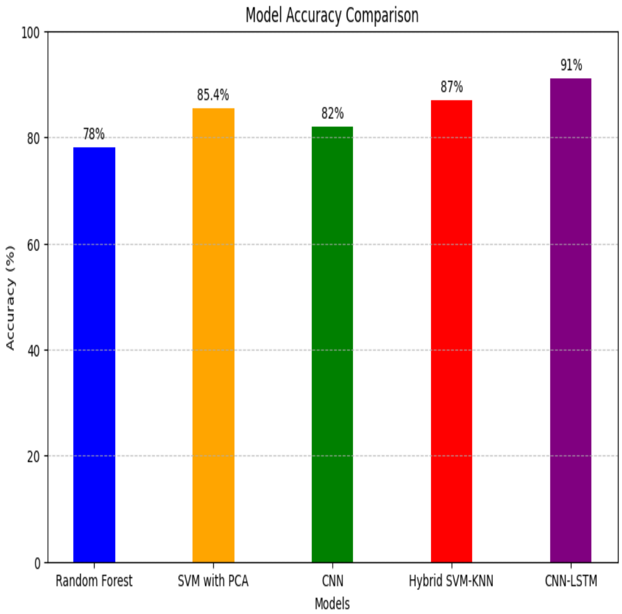



**Figure 10.** Performance Comparison of Models

## V. Conclusion and Future Scope

Action 0,1 and 2 are walking, walking upstairs, and walking downstairs respectively. Action 3 is sitting while 4 is standing and 5 is laying. Actions 0, 1, 2 and 5, We can say that we get an appropriate accuracy but for actions 3 and 4, we get the accuracy below 90% which indicates class imbalance in those 2 actions. A Multivariate Time Series consist of more than one-time dependent variable and each variable depends not only on its past values but also has some dependency on other variables. To deal with MTS, one of the most popular methods is Vector Auto Regressive Moving Average models (VARMA) that is a vector form of autoregressive integrated moving average (ARIMA) that can be used to examine the relationships among several variables in multivariate time series analysis. LSTM: Long-short term memory (LSTM)is a deep learning model; it is a kind of Recurrent Neural Network (RNN) to read the sequence dependencies. We have used CNN-LSTM. Which is made for reading long sequences of 1D time inputs with multiple features associated with that particular time frame. We have tested this model with our own self-created dataset through the use of Physics Suite Toolbox which is used as a sensor to record data further test the efficacy of the model with our own testing.

**Future Scope:**

Physics Toolbox records normally the data with the 199Hz and data recorded on which the model is trained on which is the UCI HAR Dataset contains data on 50Hz which is the premium version. As such Our data recording is fundamentally different from which the model is trained on. Therefore, the data in real time will need to be further calibrated to give accurate predictions on real time data.

## References

[1] Aggarwal, J.K., Ryoo, M.S. (2011). Human activity analysis: A review. ACM Computing Surveys, 43(3), 16.

[2] Chen, Y., Xue, Y., Zhang, H., Zhang, H. (2012). Human activity recognition based on wearable sensors. International Journal of Distributed Sensor Networks, 2012, 1-12.

[3] Preece, S.J., Goulermas, J.Y., Kenney, L.P., Howard, D., Meijer, K., Crompton, R. (2009). Activity identification using body-mounted sensors-a review of classification techniques. Physiological Measurement, 30(4), R1-R33.

[4] Bulling, A., Blanke, U., Schiele, B. (2014). A tutorial on human activity recognition using body-worn inertial sensors. ACM Computing Surveys, 46(3), 33.

[5] Khan, M.A.U., Gadekallu, T.R., Al-Jumaily, A., Shirvani, H., Hossain, M.S., Madhushankara, M. (2020). A comprehensive survey of human activity recognition using wearable sensors. Sensors, 20(3), 650.

[6] GF Harris, JJ Wertsch - Archives of physical medicine and rehabilitation, 1994 – Elsevier.

[7] Chambers, Henry G. MD; Sutherland, David H. MD. A Practical Guide to Gait Analysis. Journal of the American Academy of Orthopaedic Surgeons 10(3):p 222-231, May 2002.

[8] Hecht, Garin G. MD; Van Rysselberghe, Noelle L. MD; Young, Jeffrey L. MD; Gardner, Michael J. MD. Gait Analysis in Orthopaedic Surgery: History, Limitations, and Future Directions. Journal of the American Academy of Orthopaedic Surgeons 30(21):p e1366-e1373, November 1, 2022. | DOI: 10.5435/JAAOS-D-21-00785

[9] Baker, R. Gait analysis methods in rehabilitation. J NeuroEngineering Rehabil 3, 4 (2006). https://doi.org/10.1186/1743-0003-3-4

[10] Measurement and management of errors in quantitative gait data MH Schwartz, JP Trost, RA Wervey - Gait & posture, 2004 – Elsevier

[11] L. Lee and W. E. L. Grimson, "Gait analysis for recognition and classification," Proceedings of Fifth IEEE International Conference on Automatic Face Gesture Recognition, Washington, DC, USA, 2002, pp. 155-162, doi: 10.1109/AFGR.2002.1004148.

[12] Gait analysis MW Whittle - The soft tissues, 1993 – Elsevier

[13] A review of human activity recognition methods M Vrigkas, C Nikou, IA Kakadiaris - Frontiers in Robotics and AI, 2015 - frontiersin.org

[14] Human activity recognition: A survey C Jobanputra, J Bavishi, N Doshi - Procedia Computer Science, 2019 – Elsevier

[15] L. Xie, J. Tian, G. Ding and Q. Zhao, "Human activity recognition method based on inertial sensor and barometer," 2018 IEEE International Symposium on Inertial Sensors and Systems (INERTIAL), Lake Como, Italy, 2018, pp. 1-4

[16] N. Archana and K. Hareesh, "Real-time Human Activity Recognition Using ResNet and 3D Convolutional Neural Networks," 2021 2nd International Conference on Advances in Computing, Communication, Embedded and Secure Systems (ACCESS), Ernakulam, India, 2021, pp. 173-177.

[17] M. M. Hossain Shuvo, N. Ahmed, K. Nouduri and K. Palaniappan, "A Hybrid Approach for Human Activity Recognition with Support Vector Machine and 1D Convolutional Neural Network," 2020 IEEE Applied Imagery Pattern Recognition Workshop (AIPR), Washington DC, DC, USA, 2020, pp. 1-5

[18] A. S. Dileep, N. S. S., S. S., F. K. and S. S., "Suspicious Human Activity Recognition using 2D Pose Estimation and Convolutional Neural Network," 2022 International Conference on Wireless Communications Signal Processing and Networking (WiSPNET), Chennai, India, 2022, pp. 19-23

[19] O. D. Lara and M. A. Labrador, "A Survey on Human Activity Recognition using Wearable Sensors," in IEEE Communications Surveys & Tutorials, vol. 15, no. 3, pp. 1192-1209, Third Quarter 2013

[20] P. Nikolov, O. Boumbarov, A. Manolova, K. Tonchev and V. Poulkov, "Skeleton-Based Human Activity Recognition by Spatio-Temporal Representation and Convolutional Neural Networks with application to Cyber Physical Systems with Human in the Loop," 2018 41st International Conference on Telecommunications and Signal Processing (TSP), Athens, Greece, 2018, pp. 1-5.

[21] H. Yang, X. Wen, Y. Geng, Y. Wang, X. Wang and C. Lu, "MPJA-HAD: A Multi-Position Joint Angle Dataset for Human Activity Recognition Using Wearable Sensors," 2022 International Conference on Advanced Mechatronic Systems (ICAMechS), Toyama, Japan, 2022, pp. 178-182

[22] Anguita, D., Ghio, A., Oneto, L., Parra, X., & Reyes-Ortiz, J. L. (2013). A Public Domain Dataset for Human Activity Recognition Using Smartphones. 21st European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning.

[23] Reyes-Ortiz, J. L., Oneto, L., Samà, A., Parra, X., & Anguita, D. (2016). Transition-aware human activity recognition using smartphones. Neurocomputing, 171, 754-767

[24] Ronao, C. A., & Cho, S. B. (2016). Human activity recognition with smartphone sensors using deep learning neural networks. Expert Systems with Applications, 59, 235-244.

[25] Ronao, C. A., & Cho, S. B. (2016). Deep convolutional neural networks for human activity recognition with smartphone sensors. Neural Networks, 78, 137-147.

[26] Reyes-Ortiz, J. L., Oneto, L., Samà, A., Parra, X., & Anguita, D. (2016). Transition-aware human activity recognition using smartphones. Neurocomputing, 171, 754-767.

[27] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," in Neural Computation, vol. 9, no. 8, pp. 1735-1780, 15 Nov. 1997, doi: 10.1162/neco.1997.9.8.1735.

[28] T. J. Chin, L. Wang, K. Schindler, and D. Suter, "Extrapolating learned manifolds for human activity recognition," Proc. - Int. Conf. Image Process. ICIP, vol. 1, pp. 381–384, 2006, doi: 10.1109/ICIP.2007.4378971

[29] Y. L. Hsu, S. L. Lin, P. H. Chou, H. C. Lai, H. C. Chang, and S. C. Yang, "Application of nonparametric weighted feature extraction for an inertial-signal-based human activity recognition system," Proc. 2017 IEEE Int. Conf. Appl. Syst.Innov. Appl. Syst. Innov. Mod. Technol. ICASI 2017, no. ATmega 328, pp. 1718–1720, 2017, doi: 10.1109/ICASI.2017.7988270.

[30] J. YANG, J. CHENG, and H. LU, "Human Activity Recognition Based On The Blob Features National Laboratory of Pattern Recognition , Institute of Automation , Chinese Academy of Sciences , Beijing , China 100190," Icme 2009, pp. 358– 361, 20

[31] M. . Arshad, "Hybrid Machine Learning Techniques to detect Real Time Human Activity using UCI Dataset", EAI Endorsed Trans IoT, vol. 7, no. 26, p. e1, May 2021.