RESEARCH ARTICLE OPEN ACCESS

HUMAN ACTIVITY DETECTION USING DIVIDE AND CONQUER 1DCNN

Sehreen Mir*, Er Simranjit Kaur**
*([1]Pg Scholar, Department of Electronics and Communication Engineering, Sri Sai College of Engineering and Technology, Badhani, Pathankot
Email: Sehr7051@gmail.com)
** (Assistant Professor, Department of Electronics and Communication Engineering, Sri Sai College of Engineering and Technology, Badhani, Pathankot)

Abstract:
In pattern recognition, we use machine-learning algorithms to find regularities and patterns in data automatically. Classifying information based on knowledge or statistics learned from models and/or their representations is an example of model recognition. Knowing how the structure can be put to use is crucial. Advances have been achieved in pattern recognition in recent years. Human Activity Recognition (HAR) presents a formidable obstacle in the realm of task categorization. It requires signal processing, raw data, and excellent features to train machine learning models in order to anticipate human movement based on sensor data. It has only been recently shown that state-of-the-art outcomes may be obtained from the learning process of raw data using deep learning techniques like neural networks and recurrent neural networks. In this article, we attempt to develop a model that would forecast six groups of human actions using a division-based strategy of conquering. These activities include walking, ascending stairs, sitting, standing, and sleeping. Thirty individuals, including the dataset's participants, performed a variety of activities while wearing a smartphone strapped to their hips to collect the data used to create the dataset. This information is captured by the phone's sensors (accelerometer and gyroscope).When exploring the dataset, we split our method into two stages. Since we discovered that the data may be split in two ways, we have separated the walking, static stair climbing, and dynamic stair climbing data from the rest of the typical seated, standing, and laying down data. With 97% accuracy on test data and 98% accuracy on training data, the best model is kept for final prediction.

Keywords —:- ML, HAR, dataset, walking.

I. INTRODUCTION

The study of identifying human activities or movements using sensor data is known as Human Activity Recognition (HAR for short). Most forms of movement occur in the home, and include things like walking, chatting, standing, and sitting.They are also useful in other areas of the house, such as the kitchen and the office. Video, radar, and other forms of wireless recording technology have made it possible to remotely capture sensor data. Data may also be captured on the person using either medical equipment or a smartphone equipped with an accelerometer and gyroscope. It has traditionally been difficult and costly, needing specialized gear, to collect sensor data for event identification. Health and wellness monitoring gadgets, like smartphones, are now accessible to a wider audience and more reasonably priced.Therefore, it is more feasible to acquire sensor data from these devices, and the resulting insights into cognitive performance difficulties are more valuable. The challenge is predicting events from a sample of sensor data, often from a single or limited set of sensors. This problem is often framed as either a single job or a set of time-shared activities. Each subject will be dealing with significant variables, making recording more flexible, and there is no simple or straightforward method to match sensor data to individual human behaviors, making this a challenging task. The plan is to amass sensor data and the associated functions of known things, fit the model using this information, and then use this model to categorize the functions of previously unknown objects.Since HAR studies began, we have gained a wealth of data on performance, but different sensors, tasks, signal strengths, and durations have all been utilized in the experiments. Time series data processing methods, feature modification, selection, subtraction, and learning algorithms for feature selection, etc. Due of these variables, comparing various HAR approaches is challenging. In this study, we create a one-dimensional (1D) convolutional neural network (CNN) model for human behavior. In order to solve picture categorization issues, researchers have constructed convolutional neural network models, which are capable of learning their own internal representation of two points. Acceleration and gyroscope data collected from humans may benefit from the same method.The model is trained to identify patterns in periodic observations and to associate internal traits with various forms of activity. Let's pretend we're doing a HAR 6 class issue, where we're trying to identify people doing the things in Figure 1: walking, stair climbing, sitting, standing, and reaching. We employ a 6-class classifier, a divide-and-conquer strategy, to build a two-tiered understanding of these activities, i.e., dynamic and static tasks, first discovered using 2 classifiers or binary classifiers, and finally, determined using a 3-tiered classifier.

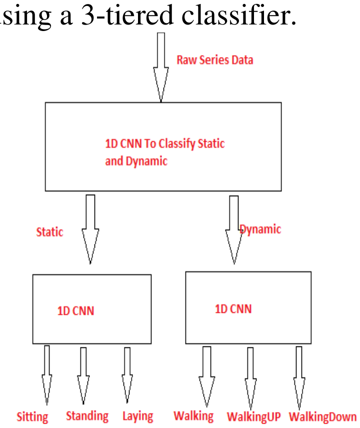

Fig. 1HAR using Divide and conquer CNN

Classifier learning is accomplished through a two-stage process: in the first stage, a binary 1D CNN model for abstract task recognition is trained to classify dynamic and static tasks; in the second phase, two 3 A 1D-like CNN models for classification of events. We can get better HAR performance by dividing multi-room problems into simpler ones. We demonstrate the effectiveness of our approach using two measured HAR data.

II. RELATED WORK

This reviews existing deep learning-based HAR and various one dimensional neural network approaches which include two-stage classifier learning techniques

A. Using logistic regression

Human activity tracking apps for smartphones have made it easier for individuals to keep track of their day-to-day routines. There have been several attempts to analyze the activities, but the results have been disappointing owing to inconsistencies in the data or defects in the various products. This research reveals a strategy for peak performance. Logistic regression may be used as a category. Smartphones, smartwatches, and other devices with built-in sensors like accelerometers and gyroscopes provide the raw data for this analysis.Since accelerometer sensors operate in three dimensions, a size reduction method must be used. Principal component analysis is useful because it allows us to distill the most essential aspects of raw data that may be used to categorize human actions. With the use of logistic approaches and open data, the authors [1] got an accuracy of 96.19 percent on the UCI-HAR dataset and 94.5 percent on the HAPT dataset.

Logical Model Trees (LTM) [2] have also been proven to be useful for activity recognition in human studies. Logistic Model Trees (LMT) include the concepts of decision tree learning with logistic regression. A decision tree model with a logistic regression function at each leaf constitutes the technique. Using cross validation, LMT identifies relevant characteristics while avoiding overfitting.

B. Using CNN

Extended short-term memory (LSTM) RNNs can learn and retain information for extremely extended periods of time. Because of this, time series analysis issues are a common application of LSTMs. Iterations in LSTMs do not include the usage of function functions. There is no fluctuation in the value stored. The information demonstrates several CNN and LSTM-based HAR applications.

Numerous implementations of HAR employing CNN and LSTM models have been shown to be effective in the literature. Using the UCI-HAR dataset, Conv2D layers for CNN, Dropout regularization, and perfect model hyper parameters in the networks of the two models, the authors of [3] present a CNN model and an LSTM model with accuracies of 99.593% and 84.71%, respectively, for six common tasks.

Combining convolutional methods with long short-term memory (LSTM) is the foundation of the deep neural networks used in the study by K. Xia et al [4].

Human activity recognition makes extensive use of deep learning techniques, which can efficiently pick out characteristics from raw data collected by sensors like gyroscopes and accelerometers. Experiments conducted and presented in [5] demonstrate that the suggested model achieves better outcomes than state-of-the-art machine learning techniques like Support Vector Machine (SVM) and k-Nearest Neighbors (KNN). We can train networks more quickly by using greater learning rates, made possible by batch normalization.

To identify six everyday activities using sensor data, Zebin et al. suggested a deep recurrent neural network (RNN) that employs a short-term model (LSTM) in [6]. The batch normalizing technique also allows more precision with fewer epochs.

The healthcare industry is another potential use of deep learning techniques. attention [7], the authors hone attention on creating a model to pinpoint the sleazy activities. We generated data for human behaviors including spitting, trash tossing, mask use, and mask non-use, and used it to train the model.CNN allows you to view picture edges, color distribution, etc., which is a huge benefit. Capturing data qualities is the capacity to extract them using kernels. Because of this, the dissemination of photos and other data types comparable to images that incorporate geographic objects takes extensive use of these networks. To do a Conv1D, the kernel moves in a single dimension. A viable choice for many 1D signal uses, particularly in situations when training data is limited. Taking into consideration current work in this field, the researchers evaluated the basic framework and ideas of 1D CNNs and significant engineering applications. [8]

C. Two-level Classifier

Decomposing a multi-class classification issue into its component parts may lead to a solution. By breaking down difficult issues into manageable chunks, we may find workable solutions and apply the findings to everyday activities. The "two-level" or "two-level" hierarchical method to characterizing activities in human understanding. Multiple investigations have shown that Human Activity Recognition (HAR) may occur in two phases.For instance, [9] suggested a hierarchical deep learning framework using the HAR model, which combines the Convolutional Neural Network (CNN) with the Bidirectional Long-Term Memory Network (BiLSTM) to learn complex representations. From the unprocessed sensor signal data, HAR may pick out the most crucial and intricate details. Using a two-tiered hierarchy, the authors extract the most useful geographic and anatomical information from a large data set at both the global and local levels. Khan et al. [10] presented a two-tiered hierarchical system as an authentication mechanism.On a more fundamental level, he classified these pursuits as either static, dynamic, or dynamic. Vectors are generated using statistical models at a higher level. Linear discriminant analysis and neural networks are then used to the resulting feature vectors in order to characterize human actions.For the purpose of categorizing photographs of artifacts on display in museums, Sandoval et al. [11] presented a two-stage image classification approach, a two-stage classifier. To begin, a convolutional network is used to segment the input picture into smaller regions, with the intention of training and classifying each region independently. The second stage involves examining the previous stage's probability vectors and deducing the final dividing order of the analyzed picture. In the field of image processing, a two-stage classifier approach was presented by Sharma et al. [12]. The diagnosis of colon cancer has been split into two stages.First, he extracted frames from the colonoscopy videos and prioritized those that showed polyps, and then he merged the data and categorised the videos as either having tumors or not.

Filios et al. [13] presented a framework for making the IoT a reality. The system is built to find individuals on the internet and uses a two-stage filtering method. The lowest levels are motion detectors and the higher ones describe various functions. The information is gathered by the smartphone's accelerometer and is a simpler model that accurately detects the user's present motion. (Times of Gratitude: sedentary, active, mobile, etc.) The user's location (e.g., at a restaurant, a grocery store, a moving automobile) and the smartphone's microphone signal are used to inform the model. Models at a higher level may take the output of models at a lower level and exercise some degree of control on those functions.

Two-stage continuous hidden Markov model (CHMM) frameworks were presented by Ronao and Cho [14]. Normal activities include walking, walking up a hill, and walking on a hill, playing mobile games or other activities at the bottom, sitting, standing, and lying down.

For human activity recognition, Zhao et al. [15] suggested a one-dimensional convolutional neural network (CNN) that uses a divide-and-conquer strategy based on categorization learning and test sharpening materials. In this approach, several 1D CNN models are trained at two different levels. In the first step, we develop a binary classifier that can divide actions into two broad categories: static and dynamic. Two separate task-recognition 1DCNN models are then used. The authors recommend information evaluation in prediction to enhance the precision of mental processes.

Similar to [15], we focus on the process of making and breaking down barriers. Classifiers are learned in two stages: first, a binary 1D CNN model is taught to distinguish between dynamic and static tasks; next, two 3A 1D-like CNN models are trained to categorize events.

.

III. PROPOSED METHOD

The methodology includes the following:

A. Data processing and transformation

Accelerometer and Gyroscope readings are taken from 30 volunteers (referred as subjects) while performing the following 6 Activities:

Walking
Walking Upstairs
Walking Downstairs
Standing
Sitting
Laying.

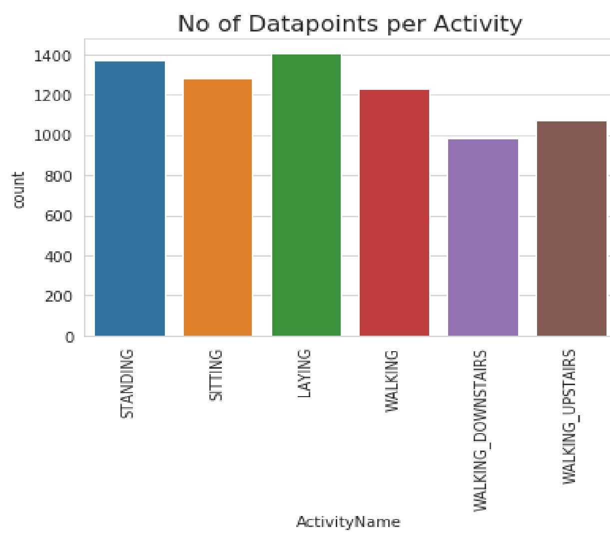

Fig. 2 Data points with activity name

Readings are divided into a window of 2.56 seconds with 50% overlapping. Accelerometer readings are divided into gravity acceleration and body acceleration readings, which has x, y and z components each. Gyroscope readings are the measure of angular velocities which has x, y and z components. Jerk signals are calculated for Body Acceleration readings. Fourier Transforms are made on the time readings to obtain frequency readings. Now, on all the base signal readings, mean, max, mad, sma, arcoefficient, energy bands, entropy etc., are calculated for each window. We get a feature vector of 561 features and these features are given in the dataset. Each window of readings is a data point of 561 features.

B. Network architecture using CNN

The CNN armature involves using Convolutional Neural Network (1DCNN) layers for point birth on input data. First step is the lading the raw dataset into memory. There are three main signals in the raw data as, total acceleration, body acceleration, and body gyroscope and each has 3 axes of data as x, y, z. therefore, there are aggregate of nine variables for each time step. Further each series of data has been partitioned into lapping windows of2.56 seconds of data or 128- time way. Thus, each row of data has 128 * 9 or 1152 elements. The affair data is defined as an integer for the class number. First these affair markers were one hot decoded so that the data will be suitable for fitting a neural network multi class bracket model. The train and test datasets will be loaded. Further scaling of the data is done using Standard Scalar object. Shape of gauged X train is(7352, 128, 9) and shape of gauged X test is(2947, 128, 9) The gauged data fits the successional model on the training dataset and evaluates it on the test dataset and returns an estimate of the model's performance. The uprooted features arealso smoothed. This is followed by a powerhouse sub caste intended to reduce over fitting of the model to the training data. Eventually, a thick completely connected sub caste is used to interpret the features uprooted by the LSTM hidden subcaste, before a final affair subcaste is used to make prognostications. In the final subcaste Softmax activation function is used since we need 6 issues as the result. The effective Adam interpretation of stochastic grade descent is used to optimize the network, and the categorical cross entropy loss function is used to calculate the loss in the training process. Once the model is fit, it's estimated on the test dataset and the accuracy of the fit model on the test dataset is returned. Accuracy on train data is given by 0.9897 and confirmation accuracy is0.9152. It's giving some good score in train as well as test but it's over fitting so important hence will use hyperactive parameter tuning Using Hyperas.

```
#best Hyper params from hyperas
eval_hyperopt_space(space, best_run)
```

```
{'Dense': 64,
 'Dropout': 0.6397045095598795,
 'batch_size': 64,
 'choiceval': 'adam',
 'filters': 32,
 'filters_1': 24,
 'kernel_size': 7,
 'kernel_size_1': 3,
 'l2': 0.07999281751224634,
 'l2_1': 0.0012673510937627475,
 'lr': 0.0011215010543928203,
 'lr_1': 0.0021517590741381726,
 'nb_epoch': 25,
 'pool_size': 3}
```

Fig. 3 Parameters of the adopted CNN model

On evaluating on train and validation data we get accuracy as follows:

Train accuracy: 0.963139281828074

Test accuracy: 0.9229725144214456

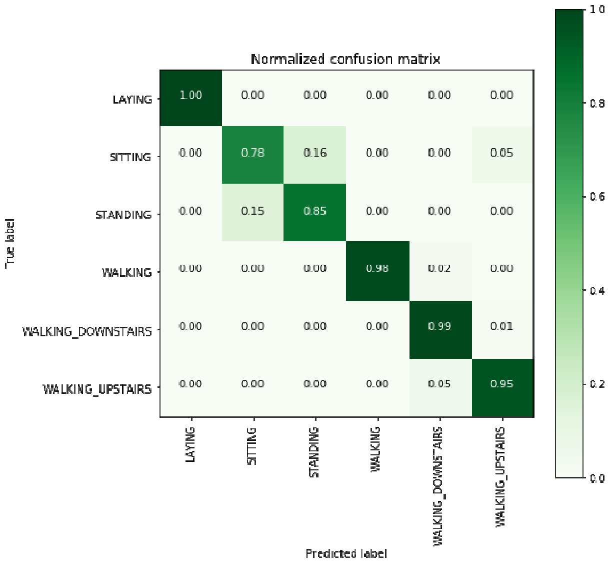

Fig.4 Normalised confusion matrix of 6-HAR Activity using CNN

.

We can observe some over fitting in the model and it is also giving some good results and error is mainly due to static activities. Therefore, we came up with some different approach to overcome this problem.

C. Network architecture using Divide and Conquer 1DCNN

1D convolutional neural network (1D CNNs) are a type of deep learning neural network that has been specifically designed to work with one-dimensional data, such as audio signals or time series data. 1D CNNs are similar to traditional CNNs, but they have a few key differences that make them well-suited for working with one-dimensional data. For example, 1D CNNs use 1D kernels (filters) that are convolved with the input data to extract features1D CNN can recognize activity such as standing, walking, jumping, and so on by receiving data from a gyroscope. Using it is advantageous. The CNNkernel, which is what other networks cannot do, is the source of CNN’s spatial features.

In Data exploration section we observed that we can divide the data into dynamic and static type so wedivided walking, walking upstairs,walking downstairs into category 0 i.e. Dynamic, sitting, standing, laying into category 1 i.e. static.

We will use 2 more classifiers separately for classifying classes of dynamic and static activities so that model can learn different features for static and dynamic activities

Our approach is two-stage activity recognition:

- By forming a model for classifying data into Static and Dynamic activities
- Classifying Static and Dynamic activities and build model for prediction on test data.

*1) **Model for classifying data into Static and Dynamic activities:*** First, we try to create a common work experience for multiple HAR units. The network architecture of our 1D CNN has two convolutional layers, a maximum pooling layer, a fully connected layer, and a softmax layer that yields the result of each of the two classes, for example; static or dynamic. The convolution is performed with a size 32 filter and a size 3 kernel. A total of 32 filters are applied to each input vector to generate feature vectors for each window size. For each eigenvector for window size and filter mode, 1-max-pooling is performed to select the largest eigenvalue. The convolutional and maximally pooled feature vectors are then used as input to combine the neural network with the output. Dropout is used to prevent the neural network from being overfitted. The softmax layer is used as the output layer of the full link layer. .Each unit in the Softmax layer calculates the probability of each static or dynamic class. The task with the highest probability is then determined as the predicted (or confirmed) task, and the task label is displayed at the end of the sequence.

2) ***Classification of Static and Dynamic activities*:**We define the base models for classification using the Keras deep learning library. The model requires a three-dimensional input with [samples, time steps, features]. We will train on 4067 samples and validate on 1560 samples for Static Activities. For Dynamic activities we train on 3285 samples and validate on 1387 samples. The output for the models will be three-element vector containing the probability of a given window belonging to each of the three activity types belonging to static activities and Dynamic respectively. These input and output dimensions are required when fitting the model, and we can extract them from the provided training dataset. The models are defined as a Sequential Keras model. We will define the models as having two 1D CNN layers, followed by a dropout layer for regularization, then a pooling layer. It is common to define CNN layers in groups of two in order to give the model a good chance of learning features from the input data. CNNs learn very quickly, so the dropout layer is intended to help slow down the learning process and hopefully result in a better final model. The pooling layer reduces the learned features to 1/4 their size, consolidating them to only the most essential elements. After the CNN and pooling, the learned features are flattened to one long vector and pass

through a fully connected layer before the output layer used to make a prediction. The fully connected layer ideally provides a buffer between the learned features and the output with the intent of interpreting the learned features before making a prediction. For this model, we will use hyper parameter tuning using hyperas. The feature maps are the number of times the input is processed or interpreted, whereas the kernel size is the number of input time steps considered as the input sequence is read or processed onto the feature maps. Transpose is used to change the dimensionality of the output, aggregating the signals by combination of sample/timestep.

```
{'Dense': 32,
 'Dense_1': 32,
 'Dropout': 0.48642317342570957,
 'choiceval': 'adam',
 'filters': 32,
 'filters_1': 32,
 'kernel_size': 7,
 'kernel_size_1': 7,
 'l2': 0.10401484931072974,
 'l2_1': 0.7228970346142163,
 'lr': 0.000772514731035696,
 'lr_1': 0.003074353392879209,
 'nb_epoch': 35,
 'pool_size': 5}
```

Fig.5 Parameters for adopted 1D CNN model.

IV. FINAL PREDICTION PIPELINE

While performing classification activity we save the best models and further use them to predict the activities. We loadkeras models and picle files for scaling data. We define function to predict activity. First we define whether an activity belongs to static or dynamic activity followed by filtering it into static and dynamic.For predicting static activities we add 4 because need to get final prediction label as output .For predicting dynamic activitieswe add1 because we need to get final prediction label as output appending final output to one list in the same sequence of input data storing asfinal_pred_val for validation data and final_pred_train for training data.

V. RESULT AND DISCUSSION

The final model used for classification of static and dynamic activities gives accuracy of 100% on train data while 0.99% accuracy on test data. This model is almost classifying data into dynamic or static correctly with very high accuracy.

```
Train_accuracy 1.0 test_accuracy 0.9989820156090939
```

Fig.6 Accuracy on test and train data using final model.

For classification of static activities using best parameters we run for 150 epochs initially and check for accuracy and loss for each epoch and approach for the number of epochs which give less overfitting with best possible results.

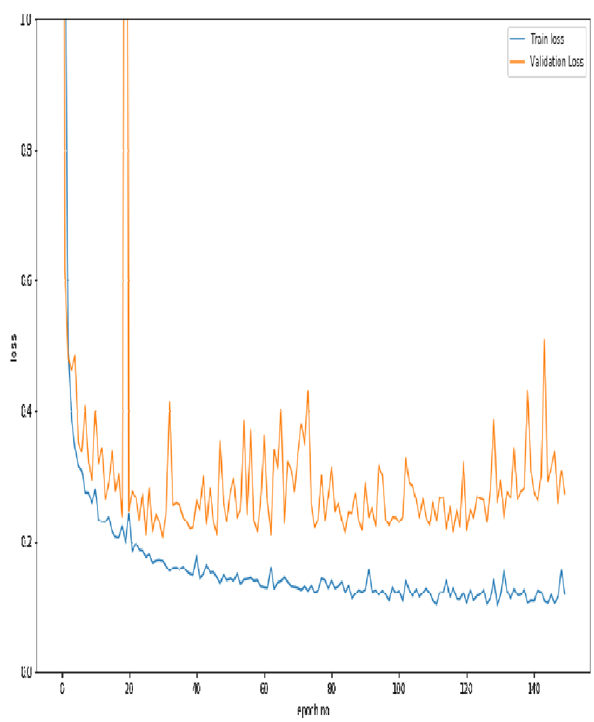

Fig.7 Loss on Static activity classifier

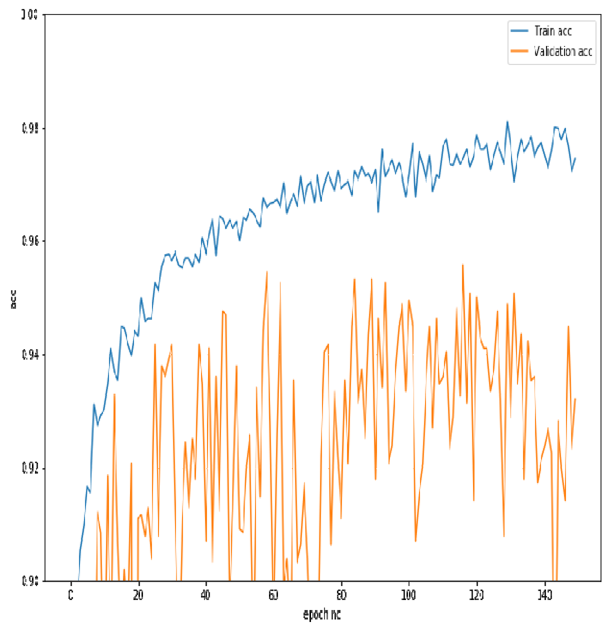

Fig.8 Accuracy of Static activity classifier.

In the graph we see around 57-59 score is giving good accuracy with less overfitting. We will run the model with epochs equal to 59.

This approach gives us better confusion metric with all data as compared models in which we approach for LSTM or CNN.

For classification of dynamic activities using best parameters we run for 70 epochs initially and check for accuracy and loss for each epoch and approach for the number of epochs which give less overfitting with best possible results.

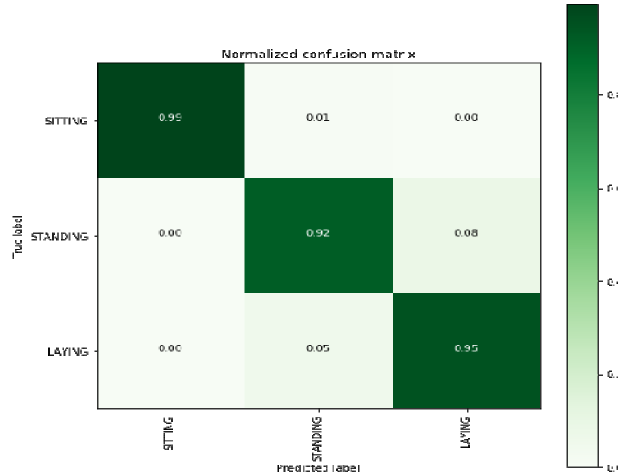

Fig.9Normalized confusion matrix for static activity classifier.

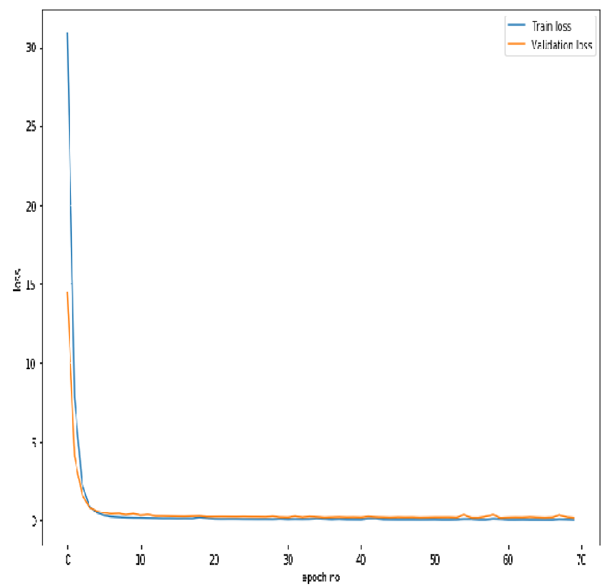

Fig.10 Loss on Dynamic activity classifier

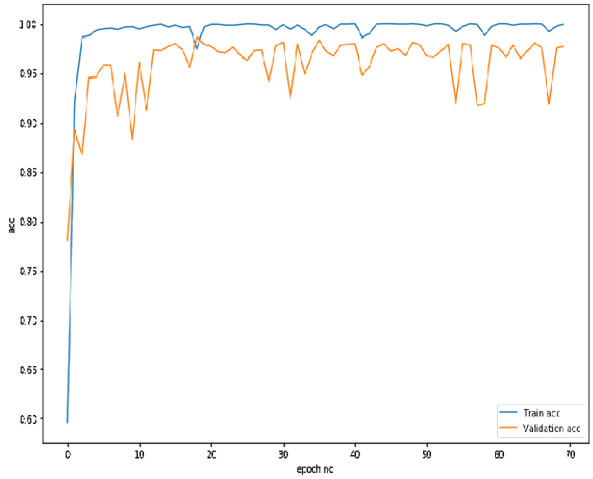

Fig.11 Accuracy on Dynamic activity classifier

In the graph we observe up to 19 epochs will give good score. We will run the model up to 19 epochs.

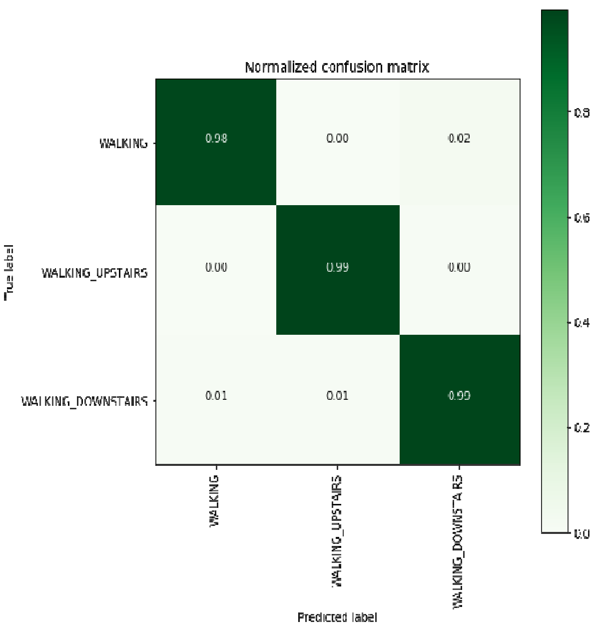

Fig.11 Normalized confusion matrix of Dynamic activity classifier

For the final pipeline predictions, the corresponding best models are saved as final models for evaluating classes and final prediction. The proposed two-stage model's complexity is high. Divide and Conquer approach with CNN is giving good result with final test accuracy of ~0.97 and train accuracy ~0.98.In general, the complexity of the two-stage model is at a disadvantage to the end-to-end model, but by replacing part of the two-stage model with other simpler models, we can reduce the complexity of the overall two-stage model. Such a strategy can be actively employed to reduce the model complexity of the two-stage models.

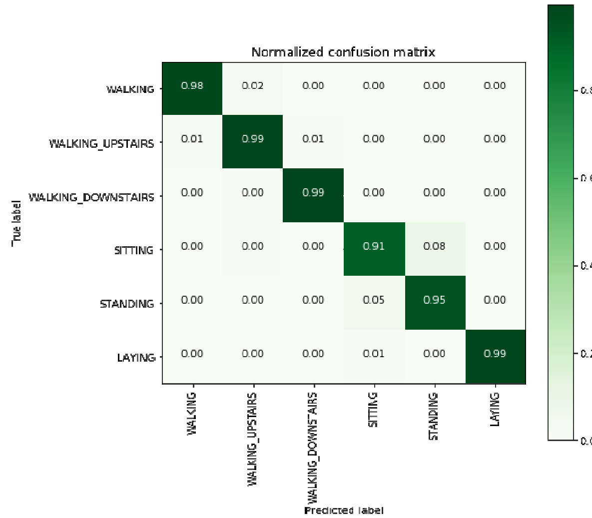

Fig.12 Normalized confusion matrix for 6-class HAR using Divide and Conquer based-1DCNN.

VI. CONCLUSIONS

We presented a divide and conquer approach for 1D CNN-based HAR for improving HAR performance.We develop a two-stage HAR method by analysing the abstract function using the confusion matrix. Our divide-and-conquer 1D CNN approach is two key elements for better HAR design. Our method is simple and practical, and it is easy to implement once the activities that fit the first phase have been identified. While using CNN architecture for a six class classifier we conclude that the accuracy of static activities on test data had some scope for improvements. Therefore by following a divide and conquer approach we could breakdown a complex classifier into two stages. At one stage, the class activities were divided for static and dynamic and a model could give a very good accuracy on its prediction. For the next stage, the classification models for classifying a three class activity were built for static and dynamic activities.

For the final prediction pipeline the best models were saved for final predictions giving approximately 97% accuracy on test data and approximately 98% on train data. Furthermore, we also explored the impact of some hyper-parameters on model performance such as the number of filters, the type of optimizers and number of epochs. Finally, the optimal hyper-parameters for the final design were selected to train the model. To sum up, compared with the methods proposed in other literatures, the Divide and conquer with CNN model shows consistent superior performance and has good generalization.

REFERENCES

[1] [1] Z. Zaki, "LOGISTIC REGRESSION BASED HUMAN ACTIVITIES RECOGNITION," JOURNAL OF MECHANICS OF CONTINUA AND MATHEMATICAL SCIENCES, vol. 15, no. 4, Apr. 2020, doi: 10.26782/jmcms.2020.04.00018.

[2] [2] A. Dehghani, T. Glatard, and E. Shihab, "Subject Cross Validation in Human Activity Recognition," Apr. 2019, [Online]. Available: http://arxiv.org/abs/1904.02666

[3] [3] C. Shiranthika, C. Shyalika, N. Premakumara, H. Samani, H.-L. Chiu, and C.-Y. Yang, "Human Activity Recognition Using CNN & LSTM Human Activity Recognition View project Reinforcement Learning based Control of Cancer Chemotherapy treatments View project Human Activity Recognition Using CNN & LSTM." [Online]. Available: https://www.researchgate.net/publication/348658435

[4] [4] K. Xia, J. Huang, and H. Wang, "LSTM-CNN Architecture for Human Activity Recognition," IEEE Access, vol. 8, pp. 56855–56866, 2020, doi: 10.1109/ACCESS.2020.2982225.

[5] [5] A. Murad and J. Y. Pyun, "Deep recurrent neural networks for human activity recognition," Sensors (Switzerland), vol. 17, no. 11, Nov. 2017, doi: 10.3390/s17112556.

[6] [6] T. Zebin, M. Sperrin, N. Peek, and A. J. Casson, "Human activity recognition from inertial sensor time-series using batch normalized deep LSTM recurrent networks," in 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), IEEE, Jul. 2018, pp. 1–4. doi: 10.1109/EMBC.2018.8513115.

[7] [7] P. Singh, I. Jindal, P. Panwar, H. Sirohi, and P. Kaushik, "Human Activity Recognition Using Deep Learning," in 2022 1st International Conference on Informatics (ICI), IEEE, Apr. 2022, pp. 75–79. doi: 10.1109/ICI53355.2022.9786890.

[8] [8] S. Kiranyaz, O. Avci, O. Abdeljaber, T. Ince, M. Gabbouj, and D. J. Inman, "1D convolutional neural networks and applications: A survey," Mech Syst Signal Process, vol. 151, p. 107398, Apr. 2021, doi: 10.1016/j.ymssp.2020.107398.

[9] [9] N. T. Hoai Thu and D. S. Han, "HiHAR: A Hierarchical Hybrid Deep Learning Architecture for Wearable Sensor-Based Human Activity Recognition," IEEE Access, vol. 9, pp. 145271–145281, 2021, doi: 10.1109/ACCESS.2021.3122298.

[10] [10] A. M. Khan, Young-Koo Lee, S. Y. Lee, and Tae-Seong Kim, "A Triaxial Accelerometer-Based Physical-Activity Recognition via Augmented-Signal Features and a Hierarchical Recognizer," IEEE Transactions on Information Technology in Biomedicine, vol. 14, no. 5, pp. 1166–1172, Sep. 2010, doi: 10.1109/TITB.2010.2051955.

[11] [11] C. Sandoval, E. Pirogova, and M. Lech, "Two-Stage Deep Learning Approach to the Classification of Fine-Art Paintings," IEEE Access, vol. 7, pp. 41770–41781, 2019, doi: 10.1109/ACCESS.2019.2907986.

[12] [12] P. Sharma, K. Bora, K. Kasugai, and B. Kumar Balabantaray, "Two Stage Classification with CNN for Colorectal Cancer Detection," Oncologie, vol. 22, no. 3, pp. 129–145, 2020, doi: 10.32604/oncologie.2020.013870.

[13] [13] G. Filios, S. Nikoletseas, C. Pavlopoulou, M. Rapti, and S. Ziegler, "Hierarchical algorithm for daily activity recognition via smartphone sensors," in 2015 IEEE 2nd World Forum on Internet of Things (WF-IoT), IEEE, Dec. 2015, pp. 381–386. doi: 10.1109/WF-IoT.2015.7389084.

[14] [14] C. A. Ronao and S.-B. Cho, "Recognizing human activities from smartphone sensors using hierarchical continuous hidden Markov models," Int J Distrib Sens Netw, vol. 13, no. 1, p. 155014771668368, Jan. 2017, doi: 10.1177/1550147716683687.

[15] H. Cho and S. Yoon, "Divide and Conquer-Based 1D CNN Human Activity Recognition Using Test Data Sharpening," Sensors, vol. 18, no. 4, p. 1055, Apr. 2018, doi: 10.3390/s18041055.

