HUMAN ACTIVITY RECOGNITION WITH SMARTPHONES

Julian J
Department of Computer Science
Christ (Deemed to be University)
Bangalore, Karnataka, India

Abstract – Human activity recognition (HAR) has a unique role in many important applications, including health-care services, ubiquitous computing, and smart buildings. Because smartphone sensors are non-intrusive, they are being used widely for identification of human activities. The signals received from smartphones are noisy, so feature engineering is performed on the data to extract more precise data. Various machine learning algorithms are used to recognize the different human activities. In this paper a comprehensive study is made on the five latest and modern technique for HAR. The techniques are Ensemble extreme learning machine, Mobile crowdsensing environment using optimized k -NN algorithm, Probabilistic First-Take-All, Convolution neural network model and Deep recurrent neural network. The major and common challenges that HAR is facing are discussed with the possible solutions. Some of the important applications for HAR are also discussed.

Keywords: Convolution neural network, Deep recurrent neural network, Ensemble extreme learning machine, Human activity recognition, k -NN algorithm, Mobile crowdsensing, Probabilistic First-Take-All.

I. INTRODUCTION

Smartphone have become a part of our daily life and the advancement in technology have made them to meet all the needs of the customer. To make the smartphones more functional manufactures add more hardware components into them. Sensors are the most important hardware components that make the smartphones more functional and powerful. Theses sensor are used to collect data from the environment and make the smartphones aware of the changes in the environment. Huge amount of data can be collected from these sensors in the smartphone. These data can be further used to find the different activities the customers are performing day to day. Accelerometer and gyroscope are the most used sensor for collecting data to analysis Human activities.

Accelerometer is one of the main sensors that is present all the smartphones. Accelerometer is used to measure the change in speed, not the speed itself. The data collected from this sensor is processed to detect sudden changes in movement. Another sensor the is present in all the smartphones is gyroscope which is used to measure orientation by using gravity. The data collected from this sensor can be processed to detect position and alignment of the device. So, these different data from these two sensors can be used to determine the activities of the person who has this smartphone [1].

Some of the latest approaches for analyzing these data to find the different human activities are:

- Ensemble extreme learning machine (ELM)
- Mobile crowdsensing environment using optimized k -NN algorithm
- Probabilistic First-Take-All (pFTA)
- Convolution neural network (CNN) model
- Deep recurrent neural network (DRNN)

1.1 Ensemble Extreme Learning Machine (ELM)

A new learning algorithm ELM for single-hidden layer feedforward neural networks (SLFNs) is used to randomly choose hidden nodes and analytically determine speed. Using ELM shows that in most cases it can learn thousands of times faster than conventional popular learning algorithms for feedforward neural networks. A comparison between ELM and support vector machines (SVM) shows than ELM may learn thousand times faster than SVM [2]. Extremely fast learning speed and good generalization performance of ELM has been applied for HAR. Ensemble learning has been added to ELM to improve the performance. Ensemble learning is to aggregate multiple base learners to boost the performance [3].

1.2 Mobile Crowdsensing Environment Using Optimized K -NN Algorithm

Individuals with computing devices such as smartphones are collectively sensing and sharing data to extract information to map and measure phenomena of common interest is called mobile crowdsensing [4]. A sample is classified by calculating the distance to the nearest training case, in k-NN the idea is extended by taking k nearest points. K-NN is a supervised learning algorithm, in which the result of a query is classified based on majority of k-NN category [5].

1.3 Probabilistic First-Take-All (pFTA)

Sensor signals of a set of human activity pattern are represented with linear projections in different subspaces. Using probability, the moments that the pattern occurs in the sequence is characterized. A set of hypotheses are defined that one of the patterns shall appear first with respect to other patterns. The hypothesis is accepted by maximum likelihood which has the largest probability by the index of the pattern that comes first. This process is repeated several times of different random projections and finally produce a compact feature vector. This feature is called as pFTA feature [6].

1.4 Convolution Neural Network (CNN) Model

When compared CNN to standard feedforward neural networks they have fewer connections and parameters and so they are easier to train. Using purely supervised learning deep convolutional neural networks with highly challenging dataset can achieve record breaking results [7]. In traditional neural network 1D vector is the input, whereas in CNN it makes a clear assumption that the input is in 3D vectors. A series of convolution and pooling is applied to the 3D vector. So, a deep CNN model will provide an effective and efficient smartphone based HAR system [8].

1.5 Deep Recurrent Neural Network (DRNN)

Deep learning is built around the idea that a deep hierarchical model is exponentially more efficient and has more functions than RNN. A deep sum-product network requires less units to represent than the sum function compared to a shallow sum-product network [9]. A newly constructed DRNN with three-axis acceleration data at time corresponding to three-dimensional input layer, and six activity classes to the six-dimensional output layer to perform high throughput activity recognition [10].

II. HUMAN ACTIVITY RECOGNITION (HAR)

HAR is a challenging and important research area with applications in many fields such as healthcare, smart environments, and homeland security. Computer vision-based techniques were widely used for activity tracking, but it requires a huge infrastructure support. Alternative and more efficient approach is by using the data collected by users' smartphones for tracking motions [11]. HAR has been explored actively based on different kinds of ambient and wearable sensors. Some of the sensors include motion, proximity, microphone, video sensors. Many HAR systems in recent years are using accelerometers to recognize a big range of daily activities such as standing, sitting, walking, running and lying down. Wearable sensors are being used in HAR, but the smartphones are the alternative to them because of the support of the diversity of sensors in it [12].

Let us now see some of the latest techniques that are being used for HAR.

2.1 Ensemble Extreme Learning Machine In HAR

In ensemble learning it tries to combine multiple base learners to boost performance. ELM algorithm randomly gives weights and biases to the input layer, which are highly unstable. Hence ELM is used with ensemble learning. A novel ensemble ELM with Gaussian random projection (GRP) is initialized for input weights of base ELMs. Data diversity is the most very important to boost performance in ensemble learning [13]. The GRP can ensure the minimal loss of information after projection. GRP is used for weight initialization which maps inputs to a random space with more diversity than uniform random values. With the normalization factor in GRP, the hidden nodes of the output are more diverse in weight initialization. The novel ELM with GRP can provide more diversity and achieving better generalization performance for ensemble learning. The raw sensor data is collected with smartphones. The statistical features are extracted on sliding window. ELM with GRP is performed on the extracted features. Majority voting scheme is used to classify human activities [2].

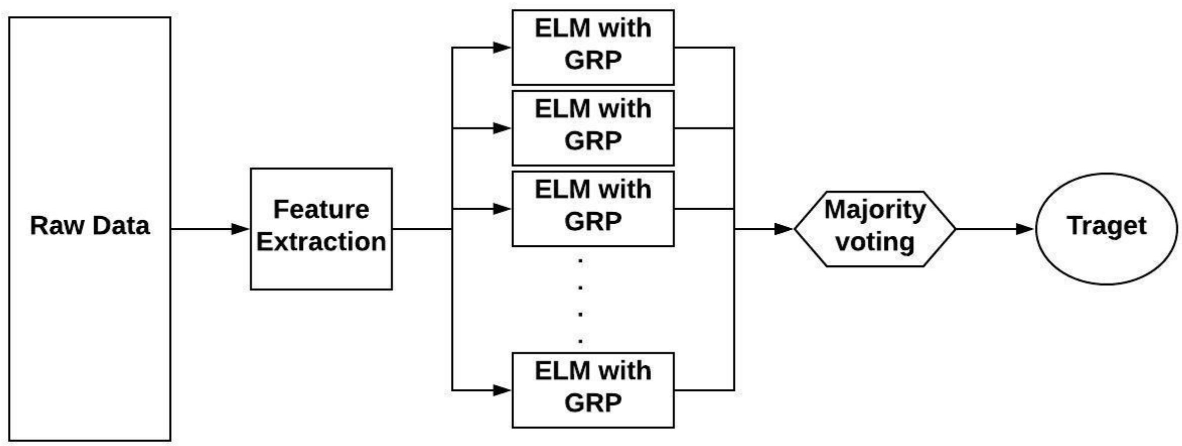

Fig. 1. Ensemble ELM for HAR

2.2 Mobile Crowdsensing Environment Using Optimized K -NN Algorithm In HAR

k-NN classifier has only one parameter. If value of k is small, noise samples will have a high impact on the classification results. If it is large, then the weight of the most similar samples will decrease, and it will need more computation time. So, we need to find the value of k that provides the minimum misclassification rate. The particle swarm optimization (PSO) is an algorithm to search for the positions that will be close to the global minimum or maximum solution. The steps involved in PSO-kNN algorithm are as follows. Firstly, data preprocessing is important because it avoids attributes of features that are larger numeric ranges which dominates the lesser numeric ranges also avoids difficulties in numerical calculations and to get higher classification accuracy [14]. In the next step all the parameters are initialized. The PSO algorithm gives k-NN classifier the value of k to find the k nearest neighbors. Next step, each testing sample will be classified for all the positions in PSO and selecting the k nearest neighbors. This is called fitness evaluation. The last step is termination criteria, when the termination criteria is reached the iteration ends or else it is proceeded with next iteration. The PSO algorithm will end when the maximum number of iterations is reached [15].

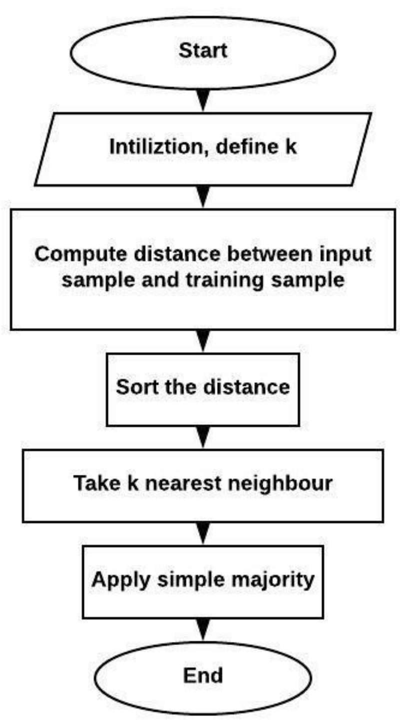

Fig. 2 k-NN algorithm

2.3 Probabilistic First-Take-All In HAR

The different properties of the pFTA algorithm will be discussed. At first the relative temporal ordering is considered rather than absolute occurring time for the latent patterns. Also, there is highly compact pFTA features vector in encoding algorithm from the original sequential data. Because of these properties, hamming distance can be used to find the similarity between to sequences, and HAR will be efficiently solved by pFTA feature vector with the nearest neighbor search [6].

2.4 Convolution Neural Network In HAR

The structure of the CNN is explained in layers. In convolution layer a Convolution operation are performed for the two vectors a ∈ Rn and b ∈ Rm where m << n. a and b are input vector and convolution filter, respectively. The activation function is performed on the output of the convolution layer to differentiate the non-linear decision boundaries of the model. The pooling layer is used to reduce dimensions of feature representation by convolution layer. The fully connected layer is responsible for applying a series of convolution and pooling for the output to flattened it to 1D vector. In the output layer, which is the last layer of CNN is used to find out the probability distribution of the predicted labels [8].

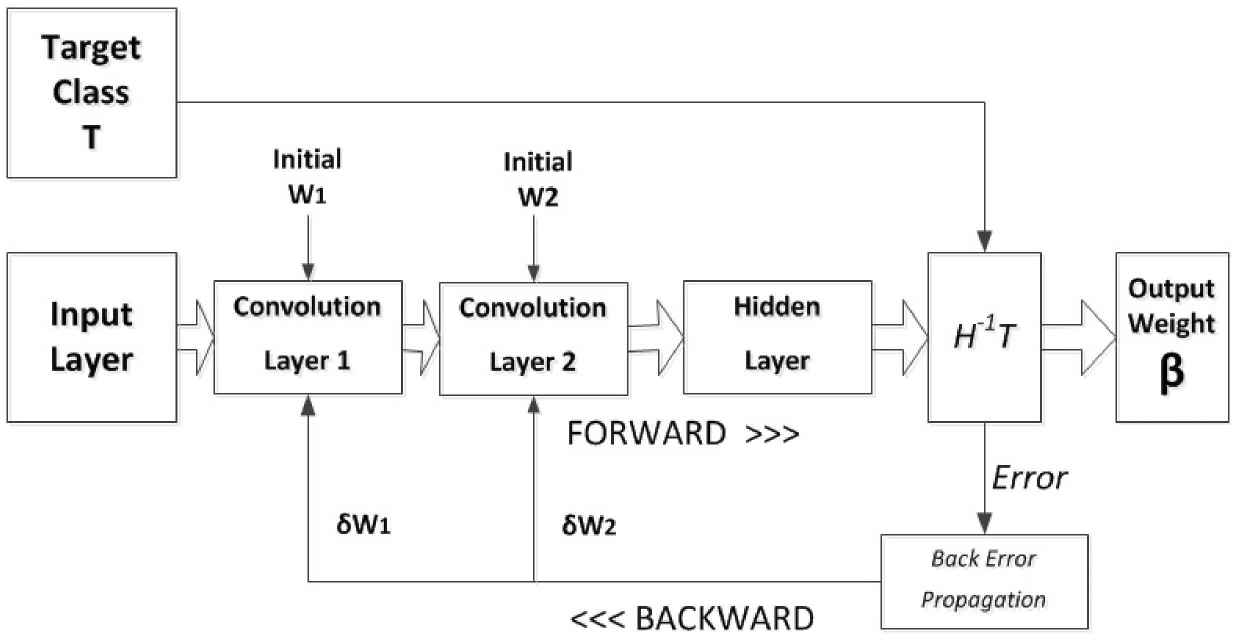

Fig. 3 CNN Architecture [16]

2.5 Deep Recurrent Neural Network In HAR

In DRNN all the unit of each internal layer is an long short-term memory (LSTM) unit. The activation function in the output layer and the error function are defined by a cross entropy function and softmax function. Backpropagation through time (BPTT) is truncated to be used to update the weights at the time of training. Each activity class corresponds to an element with largest value in the output vector. After training HAR data is executed using the forward propagation of the learned model. This shows that the training time was reduced by using the GPU for parallel processing and evaluating the throughput in the CPU of the new DRNN [10].

III. CHALLENGES IN HAR

There are many researches made on HAR in smartphones but still there are many challenges that are being faced. In this section, major and common challenges for HAR using smartphones, and the corresponding solutions will be discussed.

3.1 Subject Sensitivity

The data collected is heavily affected even by the subjects that are being selected for the training and testing. Different people have different motion patterns and even the same subject may have different patterns at different time. The accuracy decreases when the collected data is from the same subject, but on different days. The solution is to have cross-person activity recognition model which will eliminate the effects of subject sensitivity.

3.2 Location Sensitivity

The raw reading from the sensor depends on the position and orientation of the smartphone on the subject's body. There is difference in the data when the person is walking while holding the phone in his/her hand and when the phone is in his/her pocket. The solution will be by addressing another sensor, magnetometer. This sensor gives the magnetic vector along three axes in orthogonal directions. It gives a way in which the estimation can be performed before doing any motion data analysis.

3.3 Activity Complexity

Recognition model also have additional challenge because of the complexity of user activities. It is difficult to recognize the transition motion between two activities. The classifier will be confused when the user is performing multiple tasks at the same time because it is trained under one activity per segment. Hidden Markov Model (HMM) is used for smoothing the error while activity transition period.

3.4 Energy and Resource Constrains

HAR require continuous sensing of data and online updating to the classifier model which is very energy consuming. These operations also require significant computing resources. Different activities require different sampling frequency, adaptively making the choice for sampling frequency and classification features. This reduces both computing resource and energy cost. It even removes time consuming frequency feature calculation [17].

IV. APPLICATIONS FOR HAR

There are many applications that use HAR, some of the major applications that use HAR in smartphone will be discussed.

4.1 Daily Life Monitoring

All the applications in daily life monitoring are normally used to have a convenient reference for the activity logging and assisting with exercise and healthy lifestyles. Smartphone are equipped with embedded sensors such as gyroscope, GPS, accelerometer and they help track user's steps taken, stairs climbed, hours slept, quality of sleep, calorie burned, distance travelled, etc. By collecting all the data, it offers the user a complete summery on his/her lifestyle.

4.2 Personal Biometric Signature

Any two subject's motion pattern is exclusive and unique. For example, when people walk, it is almost impossible to share the exact same motion patterns. Even when there is a successful imitation, the differences are still existing because there will be difference in the motion related to bones and muscles on the human bodies. Sensors like accelerometers will be able to capture those differences. HAR provides a solution for human biometric signature with patterns in motion and gestures.

4.3 Elderly and Youth Care

One of the major goals in the current research in HAR is to develop new technologies and applications for elderly care. These applications will prevent elderly people from any harm. This will be able to detect elderly people when they are in dangerous situations. Fall detection is one such application. HAR in elders will be helpful in a proactive way, like life routine reminder (e.g., taking medicine), life monitoring in case of emergency. The youth care will also be helpful when HAR is used in children. This will help in monitoring infant's sleeping status and predicting when they are hungry.

4.4 Localization

HAR in smartphones could help in context awareness and will also be applied in localization. There are many reasons to use HAR rather than GPS for localization they are: GPS signals are very weak inside buildings and underground. GPS localization is a 2-D based positioning it does not have any information about user's altitude. GPS accuracy decreases inside cities when they have tall buildings surrounded. GPS-based localization will be confused between a movie theatre and a restaurant, which may be very close to each other. HAR related applications can be able to help by augmenting the positions with people's current activity.

4.5 Industry Manufacturing and Assisting

HAR can also assist workers in their daily work. Having a robotic exoskeleton with wearable sensors is kind of an extensions of the body which allows a worker to perform extraordinary tasks. Many other applications based on HAR include smart cameras that will be able to understand people's gestures and robot assistance in car production, etc. [17].

V. CONCLUSION

Smart phones are ubiquitous, and they are becoming more and more sophisticated. This has changed the way in which phones are being used. The landscape of people's daily life has been changing and has made way for many interesting data mining applications. HAR is the main core for all these applications. This paper presents a comprehensive survey on the latest techniques that are being used for HAR in smartphones and the challenges and applications of HAR. Technique like Ensemble ELM, mobile crowdsensing in k-NN, pFTA, CNN and DRNN are the latest and most efficient way for HAR.

References

[1] E. Bulbul, A. Cetin and I. A. Dogru, "Human Activity Recognition Using Smartphones," *2018 2nd International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT),* pp. 1-6, 2018.

[2] Z. Chen, C. Jiang and L. Xie, "A Novel Ensemble ELM for Human Activity Recognition Using Smartphone Sensors," *IEEE Transactions on Industrial Informatics,* vol. 15, no. 5, pp. 2691-2699, 2019.

[3] G.-B. Huang, Q.-Y. Zhu and C.-K. Siew, "Extreme learning machine: Theory and applications," *Neurocomputing,* vol. 70, no. 1-3, pp. 489-501, 2006.

[4] R. K. Ganti, F. Ye and H. Lei, "Mobile crowdsensing: current state and future challenges," *IEEE Communications Magazine,* vol. 49, no. 11, pp. 32-39, 2011.

[5] M. J. Islam, Q. J. Wu, M. Ahmadi and M. A. Sid-Ahmed, "Investigating the performance of naive-bayes classifiers and k-nearest neighbor classifiers," in *2007 International Conference on Convergence Information Technology (ICCIT 2007)*, Gyeongju, 2007.

[6] J. Ye, G. Qi, N. Zhuang, H. Hu and K. A. Hua, "Learning compact features for human activity recognition via probabilistic first-take-all," *IEEE transactions on pattern analysis and machine intelligence,* 2018.

[7] A. Krizhevsky, I. Sutskever and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Advances in Neural Information Processing Systems,* pp. 1097-1105, 2012.

[8] B. Almaslukh, J. A. Muhtadi and A. M. Artoli, "A robust convolutional neural network for online smartphone-based human activity recognition," *Journal of Intelligent & Fuzzy Systems,* pp. 1-12, 2018.

[9] R. Pascanu, C. Gulcehre, K. Cho and Y. Bengio, "How to construct deep recurrent neural networks," *arXiv preprint arXiv:1312.6026,* 2013.

[10] M. Inoue, S. Inoue and T. Nishida, "Deep recurrent neural network for mobile human activity recognition with high throughput," *Artificial Life and Robotics,* vol. 23, no. 2, pp. 173-185, 2018.

[11] A. Bayat, M. Pomplun and D. A. Tran, "A study on human activity recognition using accelerometer data from smartphones," *Procedia Computer Science,* vol. 34, pp. 450-457, 2014.

[12] M. M. Hassan, M. Z. Uddin, A. Mohamed and A. Almogren, "A robust human activity recognition system using smartphone sensors and deep learning," *Future Generation Computer Systems,* vol. 81, pp. 307-313, 2018.

[13] Y. Ren, L. Zhang and P. N. Suganthan, "Ensemble classification and regression-recent developments, applications and future directions," *IEEE Computational Intelligence Magazine,* vol. 11, no. 1, pp. 41-53, 2016.

[14] C. F. L. J. K. T. M. Z. Mingyuan Zhao, "Feature selection and parameter optimization for support vector machines: A new approach based on genetic algorithm with feature chromosomes,," *Expert Systems with*

Applications, vol. 38, no. 5, pp. 5197-5204, 2011.

[15] A. Tharwat, H. Mahdi, M. Elhoseny and A. E. Hassanien, "Recognizing human activity in mobile crowdsensing environment using optimized k-NN algorithm," *Expert Systems with Applications,* vol. 107, pp. 32-44, 2018.

[16] A. Budiman, M. I. Fanany and C. Basaruddin, "Distributed averaging CNN-ELM for big data," *arXiv preprint arXiv:1610.02373,* 2016.

[17] J. T. Sunny, S. M. George and J. J. Kizhakkethottam, "Applications and Challenges of Human Activity Recognition using Sensors in a Smart Environment," *International Journal for Innovative Research in Science & Technology,* vol. 2, no. 4, pp. 50-57, 2015.