Emteq Activity Recognition Challenge: Caring for Inter-user Dependency

Kohei Adachi
Department of Human Intelligence Systems
Kyushu Institute of Technology,
Kitakyushu, Fukuoka, Japan
adachi@sozolab.jp

Md Shafiqul Islam*
Department of Human Intelligence Systems
Kyushu Institute of Technology,
Kitakyushu, Fukuoka, Japan
shafiqul@sozolab.jp

Sozo Inoue
Department of Human Intelligence Systems
Kyushu Institute of Technology,
Kitakyushu, Fukuoka, Japan
sozo@sozolab.jp

ABSTRACT

Human activity recognition is a popular research area. For the last few decades, researchers are using different types of machine learning algorithms to detect activities from various types of data such as motion capture data, accelerometer data, gyroscope data etc. The researchers have used various machine-learning algorithms to detect simple as well as complex human activities. In this paper, we are focusing on "Emteq activity recognition challenge" where the task is to identify 8 types of human activities. The challenge dataset contains accelerometer, gyroscope and magnetometer data. For this purpose, we have extracted important features from the sensor data and used Random Forest classifier to detect activities. We have used one-person leave out cross validation for our proposed model. The F1 score for our proposed model on one-person leave out cross validation is around 47%. Between persons test that is one person's training data and another person's test data, the F1 score is around 30%.

CCS CONCEPTS

• Human-centered computing → Ubiquitous and mobile computing.

KEYWORDS

Activity recognition, Healthcare, Emteq Challenge

ACM Reference Format:

Kohei Adachi, Md Shafiqul Islam and Sozo Inoue. 2019. Emteq Activity Recognition Challenge: Caring for Inter-user Dependency. *In Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and the 2019 International Symposium on Wearable Computers (UbiComp/ISWC '19 Adjunct), September 9--13, 2019, London, United Kingdom.* ACM, New York, NY, USA, 6 pages. https://doi.org/

*Corresponding author.

UbiComp/ISWC '19 Adjunct, September 9–13, 2019, London, United Kingdom

ACM ISBN. https://doi.org/

1 INTRODUCTION

Human activity recognition is the task to recognize different types of activities from sensor data or video data. It is a popular research area in pervasive computing [15]. Over the past few decades, many researches have been done in this area. It has diversified application areas in real life such as robotics [6], video analysis [7], gaming, animation, surveillance [9], human computer interaction [10] etc. It is quite important for healthcare monitoring and assisted living center due to the increase in elderly population and healthcare issues [3]. The task of physical activity recognition is quite important to analyze human behavior based on their activities. In this paper, we focus on the "Emteq activity recognition challenge" [1].

Researchers are working to improve sensor-based and vison-based activity recognition [2] using different types of machine learning models. It is challenging to recognize human activities accurately due to complexity of activities, class imbalance problem, location sensitivity [8] etc. Many advanced technologies such as smart phone, wearable device, earable device have been used for monitoring the different type of human actions. These smart devices contain different types of sensors such as accelerometer, gyroscope, magnetometer etc. which provide sensor data. Over the past few years many researches have been done on sensor-based activity recognition [4]. In this system, data are collected from various sensors (e.g., accelerometer, gyroscope) and then machine learning algorithms are used for recognizing human activities.

Various Traditional machine learning [16] and deep learning [17] have been applied for recognizing human activities from sensor data. While using traditional machine learning models, the features are needed to extract manually. On the other hand, deep learning algorithms automatically extract features from raw sensor data. Convolutional Neural Network (CNN), Recursive Neural Network (RNN) model [17], Long Short Term Memory (LSTM) etc. are some deep learning models which have been used by many researchers as they provide quite good accuracy.

Alongside deep learning models, many researches have been conducted on transfer learning for human activity recognition. Transfer learning is the idea to gain knowledge from source dataset and apply to target dataset. Transferring knowledge from one user to another user using Maximum Mean Discrepancy (MMD) based transfer learning model was proposed in paper [19]. In paper [18] Unsupervised Source Selection for Activity Recognition (USSAR) algorithm was proposed. Heterogeneous Deep Convolutional Neural Network (HDCNN) to transfer knowledge from one device to another [20], Cross-domain activity recognition using Stratified Transfer Learning method [21] were also used for human activity recognition.

Besides working on sensor data, many researches have been done on motion data analysis, motion detection and recognition, which is known as human motion evaluation [12], [13]. Researchers have used Microsoft Kinect [11] to utilize 3D skeleton joint positions for detecting human motions and actions. In paper [5], multi-level hierarchical recognition framework is used on skeleton data [5]. In paper [14], linear search approach is taken for online action detection and recognition. In this paper we are using the Emteq challenge data which contains only sensor data. We have applied traditional machine learning as well as deep learning models on this dataset and found that Random Forest provides us best accuracy.

The paper is organized as follows: after introducing the overview of human activity recognition and related works in Section 1, we present Emteq challenge dataset description in Section 2. Then we introduce the proposed methodology. We explain the result analysis after that. Finally, we conclude the paper with discussion and some future works guidelines.

2 CHALLENGE DATASET DESCRIPTION

For the Emteq activity recognition challenge dataset [1], four volunteers performed activities of daily living over 3 hours. The dataset has 3 hours of labelled training data from 3 volunteers and 3 hours of unlabeled test data from a different volunteer. The dataset was collected in a simulated home environment with camera-based annotation of the volunteer activities. The simultaneous video recording of the participants was done to verify the annotations. The data was collected using a wearable device that contains Inertial Measurement Unit (IMU) sensor. Both training and test data are saved in '.csv' file. Each file contains accelerometer, gyroscope and magnetometer sensor data alongside timestamp and activity label data. The training dataset contains activity label but the testing dataset does not have any activity label. The activity type labelled "TRANSITION" is the time in which subjects 'transitioned' from one activity to the next. While training the model this "TRANSITION" activity need not to be used and in test data, this "TRANSITION" activity type is not required to be labelled.

There are 8 activities in the dataset which needs to be recognized using machine learning algorithm. The activities were performed during upright (standing stationary vs. walking) or sitting (sitting at a desk on a chair vs. sitting on a sofa). While performing these activities, the individuals may or may not be using a smartphone. So there are 4 main classes of activities (standing, walking, sitting on a chair, sitting on a sofa) and up to 8 sub-categories of activities.

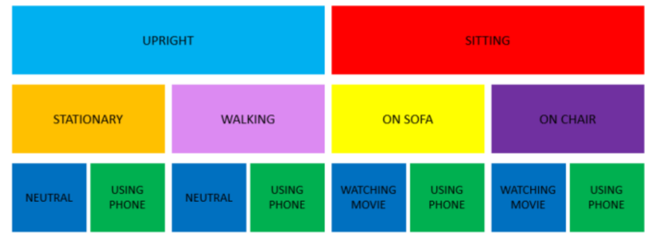

Figure 1: Activity types of different categories

In Fig. 1, all types of human activity classes based on different types of categories are shown. Some pseudo labels have been used for original activities in the dataset. The mapping of pseudo labels and activities are given below.

- 1a – upright walking
- 1b – upright walking with smartphone
- 2a – sitting on sofa watching movie
- 2b – sitting on sofa with smartphone
- 2c – sitting on chair working on laptop
- 2d – sitting on chair working on smartphone
- 3a – upright standing stationary
- 3b – upright standing with smartphone

There are around 536575 records in test dataset and 2499159 records in training dataset. There are 403593 records of activity labelled as "TRANSITION" in training dataset. Without this "TRANSITION" type, there are 2095566 records of 8 types of activities in training dataset.

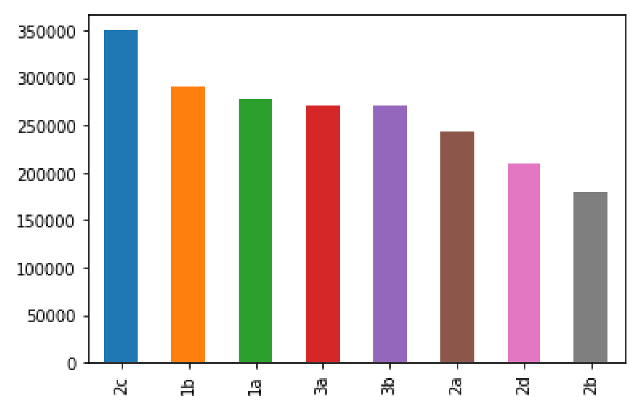

Figure 2: Activity count in training dataset

In Fig. 2, all types of activity count is shown. From the figure, it is seen that not all activity records are same. Some activities have high number of records compared to other activities.

3 PROPOSED METHODOLOGY

For human activity recognition on Emteq challenge dataset, we have used several machine learning algorithms including traditional and deep learning models. For traditional machine learning models, we have extracted features manually from dataset. On the other hand, for deep learning models it is not required to extract features manually since deep learning models automatically extract features. Among the models that we have used, Random Forest provides best F1 score. So in this section we have proposed our Random Forest classifier alongside data preprocessing and feature extraction.

3.1 Data Preprocessing

The challenge dataset contains data of 4 persons. Among these datasets, 3 person's data constitute train data and 4th person's data is the test data. The 4th person's data does not have any activity label. For training our model, we have split the 3 person's data into training and validation data. Before splitting dataset, we have preprocessed the data into several steps.

We have used leave one out cross validation for our model. There are 3 persons data in training dataset. We have done the cross validation in 2 steps. First, we have used one person's data as train data and another person's data as test data. Second, among 3 persons, we have used 2 persons data as training data and another person's data as test data.

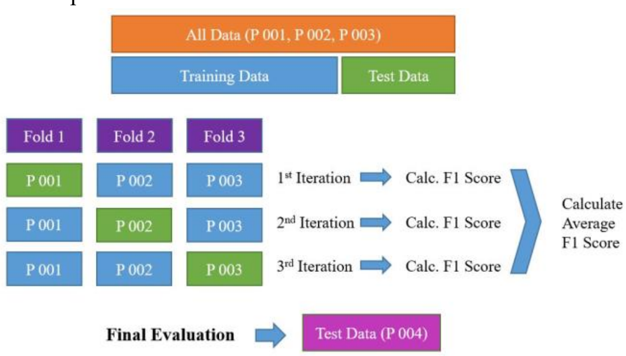

Figure 3: One person leave out cross validation

Fig. 3 represents overview of one-person leave out cross validation. Here, "P 001", "P 002", "P 003" mean the data of 1st, 2nd and 3rd person respectively. After every iteration we calculate the F1 score. After all iterations we calculate the average F1 score. Finally we evaluate the model on test data (data of person 004). There are 8 types of activities in training dataset. Each user's data file is segregated into multiple segments based on 'TRANSITION' label, which means the switching of user from one activity to another. In this way, there is no overlapping of train and test data for the model.

After splitting, missing values handling is done for both train and test data. There are several ways of handling missing values. For example, removing, imputing etc. We tried several approaches for handling missing values. We imputed the missing values with mean and std. values and found that the models accuracy is relatively lower than removing missing values. So, we have removed all missing values from train and test data.

Finally, we made a sliding window approach on both training and test data to get fixed length windows or segments from training and test data. Each window contains 100 records with 50% overlap. After finishing data preprocessing, we passed the data to our machine-learning model.

3.2 Feature Extraction

In the dataset, there are 3-axial accelerometer, gyroscope and magnetometer sensor data. Features are extracted from all of these sensor values. We have used mean, standard deviation, maximum and minimum as features for Random Forest classifier. These statistic values are calculated for each sensor's X, Y, Z values. The statistic to calculate these values are: Mean (M={Mean(X), Mean(Y), Mean(Z)}), Standard Deviation (S={SD(X), SD(Y), SD(Z)}), Maximum(Mmax={ Mmax (X), Mmax(Y), Mmax (Z)} and Minimum({Mmin (X), Mmin(Y), Mmin (Z)}). So, there are total 36 features.

3.3 Proposed Random Forest Model

We have used Random Forest classifier to detect human activities from Emteq challenge dataset and it is our proposed model. Random forest consists of a large number of individual decision trees. It creates decision trees on randomly selected data samples, gets prediction from each tree and selects the best solution by means of voting. It also provides a good indicator of the feature importance. It does not suffer from the overfitting problem since it takes the average of all the predictions, which cancels out the biases.

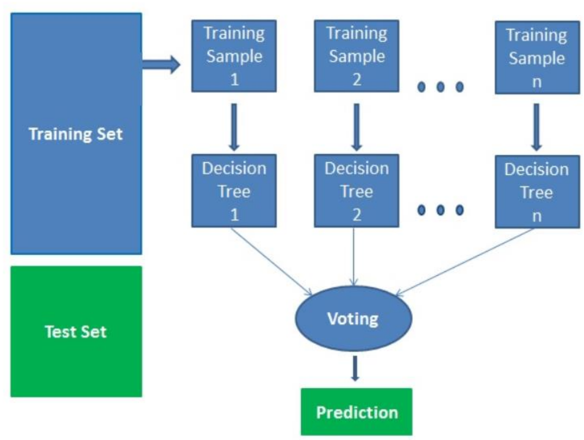

Figure 4: Random Forest Classifier

In Fig. 4, the structure of Random Forest (RnF) is shown. In RnF, N random records are picked from dataset. Then a decision tree is built based on N records. Then any number of trees are chosen for the algorithm. In case of a classification problem, each tree in the forest predicts the category to which the new record belongs. Finally, the new record is assigned to the category that wins the majority vote. For Random Forest, n_{tree} bootstrap samples are drawn from the original data. For each of the bootstrap samples, an unpruned classification tree is grown. At each node, rather than choosing the best split among all predictors, randomly sample m_{try} of the predictors are taken and the best split from among those variables are chosen [22].

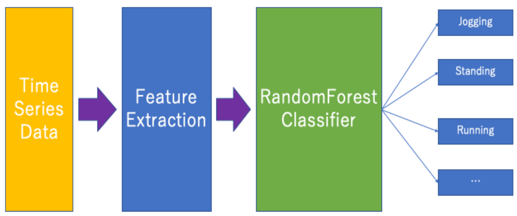

Figure 5: Architecture of proposed Random Forest model

Fig. 5 shows the overall architecture of Random Forest that we used for detecting activities. For training Random Forest classifier, we need to extract features from sensor data. We have extracted mean, standard deviation, maximum, minimum values as features. These features are extracted for each segment of training and test data. These statistic values of each segment are use as explanatory variables for Random Forest. After getting the features, we train our model. While predicting, the features of test data segments are passed to the model and activity is predicted.

4 EXPERIMENTAL RESULT

For activity recognition, we have applied several machine learning models such as KNN, Random Forest, 1D Convolutional Neural Network, Transfer Learning, Long Short Term Memory (LSTM) etc. We have used person one leave out cross validation. The experiment is done in 2 steps. First, we used 1 person's data as training data and another person's data is test data. We tested all possible combination in this way for all the machine learning models mentioned above. In the second step, we used 2 person's data as training data and 1 person's data as test data. We applied all the machine learning models for all combinations. This is our one person leave out cross validation. We did not use all persons' data at a time and then split the data into train and test data at a ratio of 7:3 or 8:2. This is done to overcome the overfitting problem.

We have used F1 score as the metric of our models performance. In one-person leave out cross validation system, after every iteration we calculate the F1 score. After finishing all iterations, we calculate the average F1 score and this is the final F1 score of the model. We have followed this procedure for all other models. F1 score is calculated using harmonic mean, not arithmetic mean. The formula of calculating F1 score:

$$F1\ Score = \frac{2\ x\ (Precision\ x\ Recall)}{(Precision + Recall)}$$

In our result analysis, we have used macro-average F1 score and weighted-average F1 score. Macro-average F1 score or macro-F1 for short, is calculated as a simple arithmetic mean of per class F1 scores. When averaging macro-F1 score, equal weights are given to each class. On the other hand, in case of weighted-average F1 score or weighted-F1 for short, the F1 score of each class is weighted by the number of samples of that class.

Table 1: F1 scores between persons

Models	Train: 1 Test: 2	Train: 1 Test: 3	Train: 2 Test: 1	Train: 2 Test: 3	Train: 3 Test: 1	Train: 3 Test: 2	Avg.
RnF (Macro F1)	32	21	28	55	12	34	30.3
RnF (Wighted F1)	33	22	30	55	12	38	**31.7**
KNN (Macro F1)	24	10	28	55	12	34	27.2
KNN (Weighted F1)	24	10	30	44	12	44	27.3
1D CNN (Macro F1)	14	17	15	43	17	38	24
1D CNN (Weighted F1)	15	18	15	44	18	43	25.5
Transfer Learning (Macro F1)	21	15	21	40	9	30	22.7
Transfer Learning (Weighted F1)	20	16	22	40	9	30	22.8
LSTM (Macro F1)	26	27	28	26	22	26	25.8
LSTM (Weighted F1)	30	28	29	27	23	27	27.3

Table 1 represents the F1 scores between persons. That is one person's data is training data and another person's data is test data. Here, train: 2 means 2nd person's data is used for training and test: 3 means 3rd person's data is used for testing. The macro-F1 score and weighted-F1 score are calculated for all the machine learning models. From the table 1, it is seen that the F1 score of Random Forest is highest among all other models. Macro-F1 score for Random Forest is around 30 % and weighted-F1 score is around 31%.

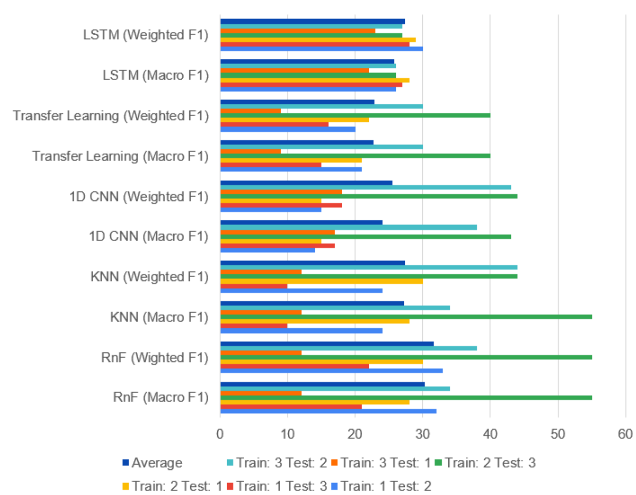

Figure 6: F1 score between persons

Fig. 6 represents the F1 score of all models in the form of bar chart. It is seen that although other models provide good F1 score in particular situation, average F1 score of Random Forest is higher than the other models. The least F1 score was found from transfer learning model when train: 3 and test: 1 data was taken.

Table 2: F1 scores for one person leave out cross validation

Models	Train:001, 002 Test:003	Train:001,003 Test:002	Train:002,003 Test:001	Average
RnF (Macro F1)	54	59	29	**47.3**
RnF (Wighted F1)	56	55	30	47
KNN (Macro F1)	49	49	22	40
KNN (Wighted F1)	41	50	22	37. 7
1D CNN (Macro F1)	33	54	14	33.7
1D CNN (Weighted F1)	34	59	15	36
Transfer Learning (Macro F1)	33	22	18	24.3
Transfer Learning (Weighted F1)	35	23	19	25. 7
LSTM (Macro F1)	31	31	29	30.3
LSTM (Weighted F1)	32	32	30	31.3

In table 2, the F1 score for one person leave out cross validation is shown. From the table 2 it is seen that the F1 score of Random Forest is highest compared to other models. Both macro-F1 and weighted-F1 score for Random Forest is around 47%.

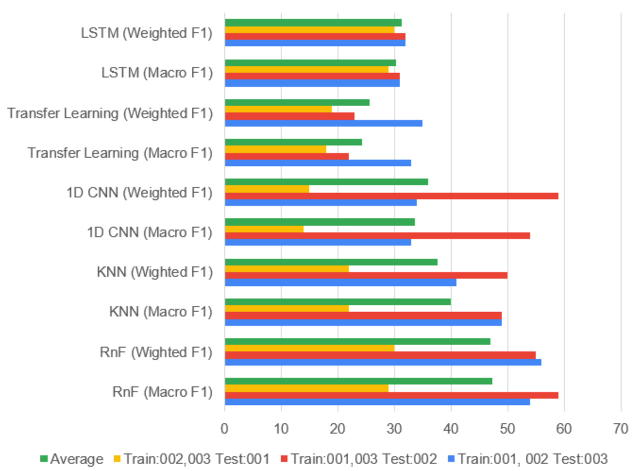

Figure 7: F1 score for one person leave out cross validation

In Fig. 7, the F1 scores of all models are shown. The average F1 score of RnF is higher than the other models. For better understanding the misclassification of the activities, confusion matrix is required. There are 3 possible scenarios since there are only 3 persons' data in the dataset. So 3 confusion matrixes are included in this paper. For the 1st confusion matrix, train data is used from person 001, 002 and test data from person 003. In the case of 2nd confusion matrix, person 001, 003 data are used for training and person 002 data for test. For the 3rd case, train data is used from person 002, 003 and test data from person 001.

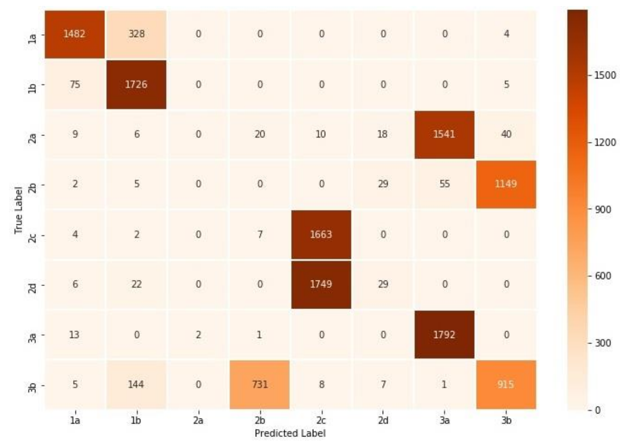

Figure 8: Confusion Matrix (Train: 001, 002 Test: 003)

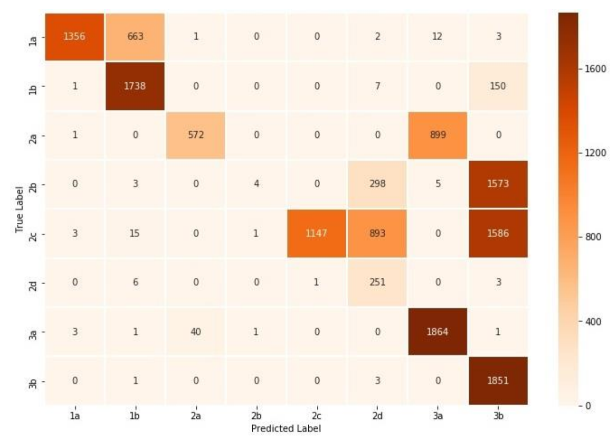

Figure 9: Confusion Matrix (Train: 001, 003 Test: 002)

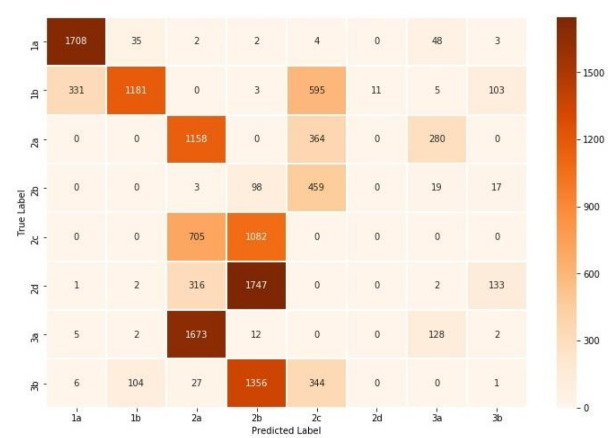

Figure 10: Confusion Matrix (Train: 002, 003 Test: 001)

Fig. 8, 9, 10 show the confusion matrixes of all possible combination of one person leave out cross validation for Random Forest classifier.

5 Discussions

From the result section, it is seen that Random Forest provides best F1 score among all other classifiers. For the one person leave out cross validation the average F1 score is 47.3% for RnF. Although it provides best F1 score still there are several misclassifications which can be seen from the confusion matrices. From the 1st confusion matrix it seen that "sitting on sofa with smartphone" and "upright standing with smartphone" causes misclassification. "Sitting on chair working on laptop" and "sitting on chair working on smartphone" also causes misclassification. In the 2nd confusion matrix, "sitting on sofa watching movie" and "upright standing stationary" causes misclassification. In the 3rd confusion matrix "sitting on sofa with smartphone" and "sitting on chair working on smartphone" causes misclassification. Besides these scenarios there are other scenarios where misclassification occurs.

The activities that cause misclassifications are almost similar. For example, "sitting on chair working with phone" and "sitting on chair working with laptop" are almost similar. It depends on how user performs these tasks. Since the activities are almost similar, so there is good chance of misclassifications. The orientation of the sensors might influence misclassification. From the confusion matrix it is seen that "upright walking" and "upright walking" activities are person independent. That means in all of those confusion matrices these activities were classified almost correctly. The rest of the activities were misclassified in different confusion matrices. So these activities are dependent on persons. Different persons perform same activity in different ways.

Regarding the F1 score, the model's performance was worst when the test data was taken from person 001. Maybe there is a significant difference between the data of person 001 and other person. This lead to low performance of the model and misclassification.

6 CONCLUSIONS AND FUTURE WORK

In this paper, we have used Random Forest classifier alongside other machine-learning algorithms for "Emteq activity recognition challenge" dataset. The Random Forest classifier provides overall best F1 score compared to other models. We extracted few features for KNN and Random Forest classifiers. Using more relevant features for activity recognition might improve the performance of our proposed Random Forest model. In future, we will use more features to improve the F1 score of our model.

REFERENCES

[1] Dataset for the emteq 2019 activity recognition challenge (www.emteq.net).
[2] T. Hossain, M. S. Islam, M. A. R. Ahad and S. Inoue, "Poster: Human Activity Recognition Using Earable Device," ACM International Joint Conference on Pervasive and Ubiquitous Computing and the International Symposium on Wearable Computers (UbiComp/ISWC '19 Adjunct), 2019.
[3] S. Abbate, M. Avvenuti, P. Corsini, J. Light, and A. Vecchio, "Monitoring of Human Movements for Fall Detection and Activities Recognition in Elderly Care Using Wireless Sensor Network: a Survey", INTECH, 2015.
[4] Z. Hajihashemi and M. Popescu, "Detection of Abnormal Sensor Patterns in Eldercare", 4th IEEE Int. Conf. on E-Health and Bioengineering, 2013.
[5] B. Su, H. Wu, M. Sheng, and C. Shen, "Accurate Hierarchical Human Actions Recognition From Kinect Skeleton Data". IEEE Access 7, 2019.
[6] L. Piyathilaka and S. Kodagoda, "Human activity recognition for domestic robots", In Field and Service Robotics. Springer, 2015.
[7] A. Jalal, Y. Kim, Y. Kim, S. Kamal, and D. Kim, "Robust human activity recognition from depth video using spatiotemporal multi-fused features", Pattern recognition 61, 2017.
[8] S. Jubil, "Applications and Challenges of Human Activity Recognition using Sensors in a Smart Environment", International Journal for Innovative Research in Science Technology, 2015.
[9] X. Li, Y. Zhang, and D. Liao, "Mining key skeleton poses with latent svm for action recognition" , Applied Computational Intelligence and Soft Computing 2017 .
[10] R. Lun and W. Zhao "A survey of applications and human motion recognition with microsoft Kinect", International Journal of Pattern Recognition and Artificial Intelligence 29, 2015.
[11] M. A. Rahman, A. M. Qamar, M. A. Ahmed, M. A. Rahman, and S. Basalamah, "Multimedia interactive therapy environment for children having physical disabilities" , In Proceedings of the 3rd ACM conference on International conference on multimedia retrieval. ACM, 2013.
[12] A. P. L. Bo, M. Hayashibe, and P. Poignet, "Joint angle estimation in rehabilitation with inertial sensors and its integration with Kinect" , Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE, 2011.
[13] Y. Chang, S. Chen, and J. Huang, "A Kinect-based system for physical rehabilitation: A pilot study for young adults with motor disabilities". Research in developmental disabilities, 2011.
[14] F. Patrona, A. Chatzitofis, D. Zarpalas, and P. Daras, "Motion analysis: Action detection, recognition and evaluation based on motion capture data" ,Pattern Recognition 76, 2018.
[15] A. Bulling, U. Blanke, and B. Schiele, "A tutorial on human activity recognition using body-worn inertial sensors," ACM Computing Surveys (CSUR), 2014.
[16] O. D. Lara and M. A. Labrador, "A survey on human activity recognition using wearable sensors," IEEE Communications Surveys & Tutorials, 2013.
[17] J. Wang, Y. Chen, S. Hao, X. Peng, and L. Hu, "Deep learning for sensor-based activity recognition: A survey," arXiv preprint arXiv:1707.03502, 17. Deep learning, Nature.
[18] J. Wang, V. W. Zheng, Y. Chen, M. Huang, "Deep Transfer Learning for Cross-domain Activity Recognition", In Proceedings of the 3rd International Conference on Crowd Science and Engineering, Singapore, 2018.
[19] D. Renjie, L. Xue, N. Lanshun, N. Jiazhen, S. Xiandong, S. Dianhui, L. Guozhong and Z. Dechen, "Empirical Study and Improvement on Deep Transfer Learning for Human Activity Recognition", Sensor, 2018.
[20] K. M. A. Hafiz, R. Nirmaly and M. Archan, "Scaling human activity recognition via deep learning-based domain adaptation" , IEEE International Conference on Pervasive Computing and Communications, Athens, 2018.
[21] W. Jindong, C. Yiqiang, H. Lisha, P. Xiaohui, and S. Y. Philip, "Stratified Transfer Learning for Cross-domain Activity Recognition" , In IEEE international conference on pervasive computing and communications (PerCom), 2018.
[22] A. Liaw and M. Wiener, "Classification and Regression by randomForest. R News", 2002.