Multi-sensor Gait Analysis for Gender Recognition

Abeer Mostafa[1][a], Toka Ossama Barghash[1], Asmaa Al-Sayed Assaf[1] and Walid Gomaa[1,2][b]

[1]*Cyber-Physical Systems Lab, Egypt Japan University of Science and Technology, Alexandria, Egypt*

[2] *Faculty of Engineering, Alexandria University, Alexandria, Egypt*

Keywords: Gender Recognition, IMU, Wavelet Transform, Supervised Learning.

Abstract: Gender recognition has been adopted recently by researchers due to its benefits in many applications such as recommendation systems and health care. The rise of using smart phones in everyday life made it very easy to have sensors like accelerometer and gyroscope in phones and other wearable devices. Here, we propose a robust method for gender recognition based on data from Inertial Measurement Unit (IMU) sensors. We explore the use of wavelet transform to extract features from the accelerometer and gyroscope signals along side with proper classifiers. Furthermore, we introduce our own collected dataset (EJUST-GINR-1) which contains samples from smart watches and IMU sensors placed at eight different parts of the human body. We investigate which sensor placements on the body best distinguish between males and females during the activity of walking. The results prove that wavelet transform can be used as a reliable feature extractor for gender recognition with high accuracy and less computations than other methods. In addition, sensors placed on the legs and waist perform better in recognizing the gender during walking than other sensors.

1 INTRODUCTION

Gender recognition has been studied widely in the last decade. Various types of data have been used to recognise the gender of a person such as images, voice signals or inertial measurements based on the motion of the person (Lu et al., 2014), (Garofalo et al., 2019) and (Zhang et al., 2017). There are many useful applications that depend on gender recognition such as speech recognition (Yuchimiuk, 2007), recommendation systems (Shepstone et al., 2013), and most importantly health care applications (Rosli et al., 2017). However, there is a huge lack of datasets and accuracy in the methods that are developed for gender recognition and the analysis of the data itself. Inertial Measurement Units (IMUs) are known to be embedded in many wearable devices which lead to useful applications. It will be convenient to recognise gender based on their readings (accelerometer, gyroscope, etc).

Datasets collected from IMU sensors are not always publicly available and most publicly available datasets don't focus on diversity of sensor placements on the human body to get the accelerometer and gyroscope signals. For these reasons, we introduce a new dataset (EJUST-GINR-1) which is collected from college students to record accelerometer and gyroscope signals from their walking activity. We record signals from smart watches and IMU sensors placed at eight different parts of the human body. We study which part of the human body effectively and uniquely identifies the gender. We run experiments on each sensor individually and also on combinations of sensors to see their effect on the classification accuracy, and in general we analyse the reliability of each body part in uniquely determining the gender of the person from the inertial movements of the corresponding body part during walking. We run experiments on a different dataset and analyse the cultural effect that can be important in changing the nature of the data. Furthermore, we propose a reliable approach to do feature extraction followed by classification to recognise the gender based on IMU readings.

There are many approaches to extract relevant features used for classification. Recently the most prominent approach is using deep neural networks. However, these methods perform well when there is a huge amount of data. This size of data is not always available when the recognition is based on data coming from sensors because the process of collecting the data and annotating it takes much time and effort. Moreover, the process may require the participation of many people and the availability of the sensors may be limited. Accordingly, we propose the use of a fea-

[a] https://orcid.org/0000-0002-8971-4311
[b] https://orcid.org/0000-0002-8518-8908

Mostafa, A., Barghash, T., Assaf, A. and Gomaa, W.
Multi-sensor Gait Analysis for Gender Recognition.
DOI: 10.5220/0009792006290636
In *Proceedings of the 17th International Conference on Informatics in Control, Automation and Robotics (ICINCO 2020)*, pages 629-636
ISBN: 978-989-758-442-8

ture extraction method that is both robust and is less dependable on the amount of available data.

Wavelet transform is a very powerful tool for the analysis and classification of signals and specifically, timeseries (Abdu-Aguye and Gomaa, 2019). However, it is unfortunately not popular within the field of data science compared with deep learning. Here, we explore the use of Wavelet transforms for feature extraction along with two classifiers: random forest and convolutional neural network (CNN) to get a reliable system for gender recognition.

The rest of the paper is organized as follows. In section 2, we do literature review of gender recognition. Then, we show some of the work that has been done using Wavelet transform specially on the signals coming from IMU sensors. In addition, we illustrate some of the work where random forest classifier has proven to be very effective on IMU signals. In section 3, we introduce our own collected dataset (EJUST-GINR-1 Dataset). Then, in section 4, we explain our methodology in detail. In section 5, we present the setup for our experiments done on our own collected dataset and the OU-ISIR Gait dataset (Ngo et al., 2014) respectively. In section 6, we discuss the main results we accomplished. Finally, we summarize our paper and show potential future work.

2 RELATED WORK

In this section, we review some of the related work and categorize them into three main categories. Research that focuses on the gender recognition problem, research techniques that apply Wavelet transform on IMU signals and research that adopts random forests or CNNs in the classification of timeseries.

2.1 Gender Recognition

Gender recognition has been adopted by researchers for many years. The variability of sensor types and applications makes it very wide and difficult research area.

The authors in (Ngo et al., 2019) presented a competition on gender and age recognition based on signals of IMU sensors placed on the waist of the person. The evaluation of models was according to performance on the OU-ISIR Gait dataset (Ngo et al., 2014). They summarize the results of gender recognition of all teams which show that most methods resulted in either a biased or inaccurate results on that dataset. The best solution used the orientation independent AE-GDI representation combined with a CNN which resulted in a classification accuracy up to 75.77%, their solution is presented in (Garofalo et al., 2019).

In the work produced by (Lu et al., 2014), the authors proposed a model to do gender recognition based on computer vision. The model tried to predict the gender of a person given a sequence of frames including arbitrary walking directions. They evaluated their model on a dataset consisting of 20 subjects (13 males and 7 females), each was captured in 4 videos. The authors reported the performance of their model which was promising for computer vision applications.

A deep learning method was used in (Zhang et al., 2017) to estimate the age and gender of a person from face images. The authors used residual networks of residual networks (RoR) as their model. The model was pre-trained on ImageNet, then it was fine-tuned on the IMDB-WIKI-101 data set for learning more complex features of face images and finally, transfer learning was done on Adience dataset. The RoR model yielded significant results compared with other deep learning techniques.

With reference to the work presented by (Jain and Kanhangad, 2016), the authors investigated solving the gender recognition problem based on data from accelerometer and gyroscope sensors which are integrated in a smart phone. The authors explored using multi-level local pattern (MLP) and local binary pattern (LBP) in feature extraction. For classification, the authors tried support vector machine (SVM) and aggregate bootstrapping (bagging). All these models were evaluated on a 252 gait dataset collected from 42 subjects and yield accuracy up to 77.45% by MLP and bagging.

2.2 Wavelet Transform on IMU Signals

Researchers have adopted the use of wavelet transform for the analysis of signals. Here, we present some examples where wavelet transform has proven to be a very robust approach.

The authors in (Abdu-Aguye and Gomaa, 2019) used wavelet transform followed by adaptive pooling to do feature extraction for human activity recognition based on accelerometer and gyroscope signals. Their approach was evaluated on seven different activity recognition datasets and yielded significant results.

In (Zhenyu He, 2010), wavelet transform was applied on 3D acceleration signals then applying an autoregressive model on the decomposed signal. The outcome coefficients were then used as the feature vector which was fed to a support vector machine classifier to distinguish between the different human activities. This model was tested on four different human activities and resulted in high accuracy classification (95.45%) which clearly shows that the approach can successfully be used in human activity recognition based on acceleration signals.

In (Assam and Seidl, 2014), the authors applied wavelet transform alongside with vector quantization and Hidden Markov Model (HMM) on sensory data from Android smart phones. The model aimed to extract the spectral features of accelerometer sensor signals by performing multi-resolution wavelet transform and using them for human activity recognition. The model result was very significant as it reached classification accuracy up to 96.15% on six human activities.

From the previous works, it seems evident that wavelet transform performs well in the analysis of accelerometer and gyroscope signals. However, it was used only for activity recognition. In our work, we explore using wavelet transform for gender recognition and seeing if the signal decomposition still able to distinguish between male signals and female signals fixing a particular activity, which is ‘walking’ in the current work. We also investigate which body part(s) are the best in uniquely differentiating gender based on inertial signals. Finally, we investigate the cultural impact and hypothesize that the effectiveness of inertial signals in gender recognition may be dependent on the culture from where the subjects come from.

2.3 Random Forests

The work presented in (Mehrang et al., 2018) uses random forests as the main classifier to recognize human activity based on triaxial acceleration signals. The system achieved accuracy of $89.6 \pm 3.9\%$ with a forest of size 64 trees.

The authors in (Feng et al., 2015) designed an ensemble learning algorithm that integrates many individual random forest classifiers. Their model was evaluated on a dataset consisting of 19 different physical activities and reached accuracy up to 93.44% with a small training time compared to other classification methods.

The authors in (Casale et al., 2011) also adopted the technique of using random forests classifier to recognize human activities based on acceleration signals. They obtained a high classification accuracy which was up to 94%.

From these results, we can conclude that the random forests classifier can be efficiently applied on accelerometer and gyroscope signals to perform gender recognition.

3 EJUST-GINR-1 DATASET

The dataset was collected using six IMU units and two smart watches. Each IMU unit is a MetaMotionR (MMR) sensor, which is a wearable device that provides real-time and continuous motion tracking (MBIENTLAB, 2018). We record the readings of the following sensors: Accelerometer, Gyroscope, Magnetometer and Pressure. The internal components of each MetaMotionR sensor is shown in Figure 1.

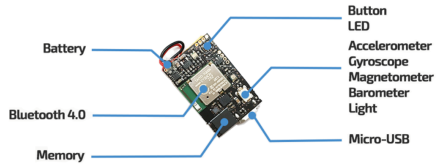

Figure 1: The components of each MetaMotionR unit (MBIENTLAB, 2018).

However, within the scope of this work, we only use the accelerometer and gyroscope. The MetaMotionR sensor specifications are illustrated in Table 1. The sensors were placed at six positions: right upper arm (RUA), left upper arm (LUA), right cube (RC), left cube (LC), waist and back, along side with two Apple watches (LH) and (RH) as shown in Figure 2.

The two watches model is series-1, integrated in them Apple S1 computer which is described as a System in Package (SiP). The SiP includes the two sensors we need: accelerometer and gyroscope. Figure 3 shows the Apple watch series-1 and the sensor axes.

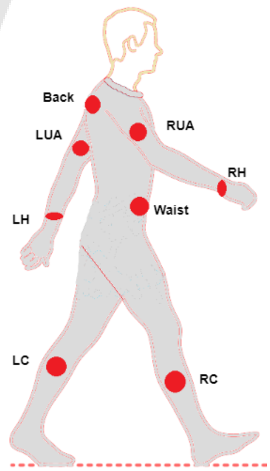

Figure 2: MetaMotionR units and smart watches placements on the human body as indicated by the red spots.

The sensors were synchronized together alongside with the smart watches to generate gyroscope and

Table 1: Sensor specification of the MetaMotionR unit.

Description	Ranges	Resolution	Sample Rate
Gyroscope	±125, ±250, ±500, ±1000, ±2000 deg/s	16 bit	0.001Hz, 100Hz stream, 800Hz log
Accelerometer	±2, ±4, ±8, ±16g	16 bit	0.001Hz, 100Hz stream, 800Hz log

Figure 3: Apple watch series-1 used for recording accelerometer and gyroscope signals.

accelerometer readings with frequency equals to 50 Hertz.

The subjects who participated in collecting this dataset are all volunteer students at our university (both postgraduate and undergraduate). The total number of data samples and subjects information are summarized in Table 2. Gait procedure: Each subject walked alone on a straight ground for 4 sessions, each session lasted for 5 minutes, totalling over than one million sensor readings. The process was standardized among all subjects. Males and females were wearing trousers in order not to change the readings of the sensors placed on the subject's legs. Participants were asked to walk naturally in the same way they walk every day. The dataset is available upon request.

4 METHODOLOGY

4.1 Feature Extraction

In our domain, we are dealing with timeseries which are the signals coming from IMU sensors. In order to analyse the timeseries, we would like to know which frequencies are present in the signal. Fourier transform is a famous method to do that. However,

Table 2: Total number of samples and subjects information of EJUST-GINR-1 dataset.

Attribute	Value
Total Number of Samples	5292
Number of Females	10
Number of Males	10
Age Range	19-33
Height Range	146cm-187cm
Weight Range	56kg-130kg

Fourier transform doesn't give any information about time (it has a high resolution in the frequency-domain but zero resolution in the time-domain) as explained in Figure 4. For that reason, scientists proposed the use of Short-Time Fourier transform. The main problem with this approach is that we face the same limits of Fourier Transform known as the uncertainty principle. The smaller we make the size of the window the more we will know about where a frequency has occurred in the signal, but less about the frequency value itself (Taspinar, 2018).

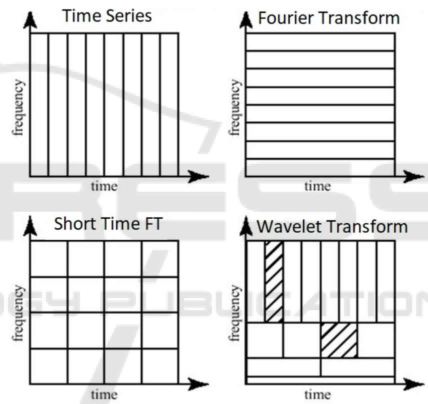

Figure 4: An overview of time and frequency resolutions of various transformations. The size and orientations of the block gives an indication of the resolution size (Taspinar, 2018).

To solve this problem, Wavelet transforms are used as they provide high resolution in frequency domain and also in time domain. This means we can know which frequencies are present in a signal and also at what time these frequencies have occurred.

Unlike Fourier transform, the Wavelet transform represents a signal as a decomposition of some functions called *Wavelets*. Each wavelet is at a different scale. The difference between wavelets and sine waves is that wavelets are positioned in time. Mathematically, the Wavelet transform is described by equation (1):

$$W(a,b) = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{\infty} x(t)\psi(\frac{t-b}{a})dt \qquad (1)$$

Here, $W(a,b)$ is the wavelet coefficient, a is the scaling variable, b is the shifting variable and $\Psi(t)$ is called the mother wavelet. Practically, these coefficients are calculated from the correlation between the signal at a time instance t and a wavelet shifted to the same time instance. The accuracy of the signal representation depends on how good we choose the wavelet. There are various families of wavelets in the literature such as Haar, Complex Gaussian wavelets, Morlet, etc.

In this work, we use the Morlet wavelet which is described by equation (2):

$$\psi(t) = \exp(\frac{-t^2}{2})\cos(5t) \tag{2}$$

Now, the coefficients obtained from Morlet wavelet transform are considered a good descriptor of the original signals. We then take these coefficients as the feature vector of our dataset and feed them to the classifier.

4.2 Classification

There are many methods in supervised machine learning that solve classification problems. Some of them do both feature extraction and classification. However, we set the feature vector as the outcome from the wavelet transform, then we consider two different classifiers: Random forest and Convolutional Neural Network (CNN).

A. Random Forest. As shown in our literature review, random forests is one of the most popular and efficient techniques as it is based on ensemble learning. Ensemble learning means that the model we use makes predictions based on many different individual mini-models. Ensemble learning is done in two ways, either bagging or boosting. Bagging means that the individual models are trained in parallel and each one uses a subset of the dataset. A random forest uses the decision trees as its individual models with bagging. In addition it does random feature selection to be able to improve the classification accuracy by reducing the prediction variance. In this work, we use a random forest classifier with its specification shown in section 5.

B. Convolutional Neural Network. CNN is a very famous approach to perform classification on multi-dimensional data. Although CNNs are considered a deep learning method, we don't use a deep network in this work for three reasons. Firstly, we don't need the network to learn more complex features as we already have our feature vector from the Wavelet transform. Secondly, a deep convolutional network will require much computations and we want to minimize the computations on the dataset as possible as the current research can be later used for online implementation on wearable devices. Finally, we would need a lot more data in deep learning to prevent overfitting. For these reasons, we use a shallow network as a classifier with its specifications described in section 5.

5 EXPERIMENTAL SETUP

To investigate the effectiveness of our methodology, we evaluate the performance on two datasets: our own collected EJUST-GINR-1 dataset and The OU-ISIR Gait Dataset (Ngo et al., 2014). Both datasets include accelerometer and gyroscope signals from IMU sensors collected for gait analysis and identification of human attributes.

5.1 Datasets Considered

5.1.1 Experiments Setup on EJUST-GINR-1 Dataset

Using EJUST-GINR-1 dataset, introduced in section 3, we ran some experiments using our model. Firstly, we use signals coming from each of the eight sensors individually to know which part(s) of the human body best distinguishes between males and females during walking. Then, we use many combinations of sensors to see if this will have an impact on the predictive performance.

The dataset was split into fixed-size samples. Each sample corresponds to a 5-second signal with its label indicating whether the subject is a male or a female. The sampling rate was 50 Hz so, each sample had 250 sensor readings with each reading consisting of six components: Accelerometer-X, Accelerometer-Y, Accelerometer-Z, Gyroscope-X, Gyroscope-Y, and Gyroscope-Z.

5.1.2 Experiments Setup on the OU-ISIR Gait Dataset

The OU-ISIR Gait dataset was collected at Osaka University to help research in the area of human identification based on gait analysis (Ngo et al., 2014). We had the permission to use the dataset in our research from the dataset administrator with a signed agreement from EJUST University to Osaka University.

The dataset was collected using three IMU sensors and a smart phone, all located around the waist of the subject. The dataset included three gait styles:

level walk, up slope and down slope. The dataset includes readings from IMUs similar to the sensors we used to collect our dataset. Each unit generates six-dimensional data: Accelerometer-X, Accelerometer-Y, Accelerometer-Z, Gyroscope-X, Gyroscope-Y, and Gyroscope-Z. The dataset was aggregated to different versions to satisfy many protocols for different research goals. We used two versions of the dataset. Firstly, the one with the largest number of subjects (total 744 subjects). This version has readings from only the sensor placed on the centre of the waist for level walk activity with sampling rate equals to 100 Hz, each subject has two signals of level walk. The total number of samples and subjects information of level walk dataset version are summarized in Table 3. The second version has a less number of subjects (total 495 subjects) and contains two signals of level walk, one for slope up and one for slope down for each subject shown in Table 4.

The length of the signals is not included in the dataset description. However, from the data itself we can conclude that each signal has a length of only few seconds.

Table 3: Total number of samples and subjects information of level walk only OU-ISIR Gait dataset version.

Attribute	Value
Total Number of Samples	1488
Total Number of Subjects	744
Age Range	2-78

Table 4: Total number of samples and subjects information of level walk, slope up and slope down OU-ISIR Gait dataset version.

Attribute	Value
Total Number of Samples	1980
Total Number of Subjects	495
Age Range	2-78

5.2 Model Specification

Our model for the EJUST-GINR-1 dataset is described as follows. First, we feed the fixed length signals to the feature extractor, which applies Morlet wavelet transform. We tried different scales of decomposition and selected the range of scales that minimizes the computations and gives a high accuracy. The scales range from 1 to 64 gave the best results. We take all the coefficients obtained from the signal decomposition and consider them the feature vector then feed it to the classifier.

In the random forest classifier, we set the number of decision trees in the forest to 100 trees and use the Gini index to measure the quality of the split. We use bootstrap aggregation to randomly select subsets of the whole dataset and also random subsets of the features. We run each experiment 10 times and report the average accuracy.

In the CNN classifier, we use a shallow network consisting of 5 convolutional layers, each followed by a max pooling layer, and at the end, one fully connected layer then the final layer that produces the binary classification output. The activation function at all layers are ReLU (Rectified Linear Unit) except at the output layer, we use softmax as activation to produce the output scores. All the code components are available on a GitHub repository and available upon request.

6 RESULTS AND DISCUSSION

In this section, we include the results obtained from the experimental evaluation of our methodology. We consider the classification accuracy as the evaluation criteria to our model.

6.1 Evaluation on the EJUST-GINR-1 Dataset

The results of our approach using wavelet transform to extract features followed by random forest classifier evaluated on the EJUST-GINR-1 dataset are shown in Table 5. In Table 5, we refer to the sensors placed in the left upper arm, right upper arm, left cube, right cube, left hand and write hand as LUA, RUA, LC, RC, LH and RH respectively along side with back and waist sensors without name abbreviations. The diagonal elements of the table represent the classification accuracy obtained by testing on the signals of each sensor individually. The off-diagonal elements represent the classification accuracy obtained by combining the data of each sensor with the data of each of the other seven sensors in order. For the overall performance, all results of individual sensors lie between 85.96% for the back sensor and 95.85% for the left cube sensor. In general, we can conclude from the results that the sensors located at lower part of the body (right cube, left cube and waist) classify gender by significantly higher accuracy than the sensors located at the upper part of the body.

Evaluating the model on 12-dimensional data by combining the accelerometer and gyroscope readings of each two sensors boosted the performance of individual sensors. As illustrated in Table 5, the accuracy obtained from The left upper arm sensor was 89.06%, and from right upper arm 86.72% but, when the two sensors combined together, the accuracy reached 94.26%. The performance also was boosted for many

Table 5: Accuracy obtained by combinations of sensors.

Sensor	LUA	RUA	LC	RC	Back	Waist	LH	RH
LUA	89.0578%	94.2647%	94.3103%	94.7267%	89.4118%	92.6782%	92.9979%	93.1703%
RUA	94.2647%	86.7155%	95.9459%	94.6705%	88.8400%	95.0291%	93.8272%	90.6826%
LC	94.3103%	95.9459%	95.8451%	94.7767%	93.8143%	96.7388%	96.5509%	94.4693%
RC	94.7267%	94.6705%	94.7767%	93.4627%	95.0348%	96.0169%	96.5698%	93.6639%
Back	89.4118%	88.8400%	93.8143%	95.0348%	85.9539%	94.8727%	91.0545%	90.8710%
Waist	92.6782%	95.0291%	96.7388%	96.0169%	94.8727%	95.7143%	96.6539%	95.3160%
LH	92.9979%	93.8272%	96.5509%	96.5698%	91.0545%	96.6539%	89.0566%	91.2606%
RH	93.1703%	90.6826%	94.4693%	93.6639%	90.8710%	95.3160%	91.2606%	88.9023%

Table 6: Our approach compared to previous approaches on the OU-ISIR dataset (Garofalo et al., 2019).

Method	Classification Accuracy
AutoWeka 2.0	58.25%
HMM	41.75%
TCN	60.31%
TCN + Orientation Independent	67.01%
CNN + AE-GDI	75.77%
Ensembel	64.43%
Wavelet + CNN	70.85%
Wavelet + random forest	68.27%

other combinations. The highest accuracy was obtained by combining the left cube sensor with the waist sensor with accuracy reaching 96.74%.

6.2 Evaluation on the OU-ISIR Gait Dataset

We evaluated our methodology on both versions of the OU-ISIR Gait dataset (Ngo et al., 2014) described in section 5 using wavelet transform as a feature extractor then trying both classifiers random forest and CNN. Both versions of the dataset give relatively lower results than EJUST-GINR-1 dataset. Our model evaluated on the first version with the highest number of subjects (744 subjects) reached 74.73% accuracy using CNN classifier specified in section 5, and 69.03% accuracy using the random forest classifier.

In the second version of the dataset, which includes less number of subjects (495 subjects) but more gait styles, we obtain accuracy up to 70.85% using CNN and 68.27% using random forest. We compare our results on the second version of the dataset to other approaches, all summarized in Table 6. As shown in Table 6, using wavelet transform as a feature extractor outperforms most other approaches. We should also mention that the computation power and time needed for applying wavelet transform is significantly less than any deep learning technique.

7 CONCLUSION AND FUTURE WORK

In this work, we proposed a reliable model for gender recognition based on inertial data of accelerometer and gyroscope signals streamed from wearable IMU units. We use wavelet transform as our feature descriptor. We also proposed a new gait dataset EJUST-GINR-1 collected from smart watches and IMU sensors placed at eight different parts of the human body. We investigated which body locations, in terms of sensors placements, best distinguish between males and females in walking. We evaluated our model on two datasets and showed the results of each dataset. Our approach gives very promising results which shows that wavelet transform can efficiently be used to extract features for gender recognition along with potentially diverse set of classifiers.

In the future, we intend to expand our approach to make age predictions, as both a classification problem (age level) as well as regression problem (estimate the exact age), and expand our dataset to include more age ranges and more activities. We would like to also investigate which other activities/actions can reliably differentiate gender using inertial sensors. These actions can include brushing teeth, sitting, standing, etc. We may also investigate the use of wavelet transform to analyse electroencephalogram (EEG) signals to do gender recognition based on brain signals analysis.

ACKNOWLEDGMENTS

This work is funded by the Information Technology Industry Development Agency (ITIDA), Information Technology Academia Collaboration (ITAC) Program, Egypt – Grant Number (PRP2019.R26.1 - A Robust Wearable Activity Recognition System based on IMU Signals).

We also would like to thank all the students who participated in collecting the dataset.

REFERENCES

Abdu-Aguye, M. G. and Gomaa, W. (2019). Competitive feature extraction for activity recognition based on wavelet transforms and adaptive pooling. In *2019 International Joint Conference on Neural Networks (IJCNN)*, pages 1–8.

Assam, R. and Seidl, T. (2014). Activity recognition from sensors using dyadic wavelets and hidden markov model. In *2014 IEEE 10th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*, pages 442–448.

Casale, P., Pujol, O., and Radeva, P. (2011). Human activity recognition from accelerometer data using a wearable device. In *Iberian Conference on Pattern Recognition and Image Analysis*, pages 289–296. Springer.

Feng, Z., Mo, L., and Li, M. (2015). A random forest-based ensemble method for activity recognition. In *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 5074–5077. IEEE.

Garofalo, G., Argones Rúa, E., Preuveneers, D., Joosen, W., et al. (2019). A systematic comparison of age and gender prediction on imu sensor-based gait traces. *Sensors*, 19(13):2945.

Jain, A. and Kanhangad, V. (2016). Investigating gender recognition in smartphones using accelerometer and gyroscope sensor readings. In *2016 International Conference on Computational Techniques in Information and Communication Technologies (ICCTICT)*, pages 597–602.

Lu, J., Wang, G., and Moulin, P. (2014). Human identity and gender recognition from gait sequences with arbitrary walking directions. *IEEE Transactions on Information Forensics and Security*, 9(1):51–61.

MBIENTLAB (2018). https://mbientlab.com.

Mehrang, S., Pietilä, J., and Korhonen, I. (2018). An activity recognition framework deploying the random forest classifier and a single optical heart rate monitoring and triaxial accelerometer wrist-band. *Sensors*, 18(2):613.

Ngo, T. T., Ahad, M. A. R., Antar, A. D., Ahmed, M., Muramatsu, D., Makihara, Y., Yagi, Y., Inoue, S., Hossain, T., and Hattori, Y. (2019). Ou-isir wearable sensor-based gait challenge: Age and gender. In *Proceedings of the 12th IAPR International Conference on Biometrics, ICB*.

Ngo, T. T., Makihara, Y., Nagahara, H., Mukaigawa, Y., and Yagi, Y. (2014). The largest inertial sensor-based gait database and performance evaluation of gait-based personal authentication. *Pattern Recognition*, 47(1):228–237.

Rosli, N. A. I. M., Rahman, M. A. A., Balakrishnan, M., Komeda, T., Mazlan, S. A., and Zamzuri, H. (2017). Improved gender recognition during stepping activity for rehab application using the combinatorial fusion approach of emg and hrv. *Applied Sciences*, 7(4):348.

Shepstone, S. E., Tan, Z.-H., and Jensen, S. H. (2013). Demographic recommendation by means of group profile elicitation using speaker age and gender recognition. In *INTERSPEECH*, pages 2827–2831.

Taspinar, A. (2018). A guide for using the wavelet transform in machine learning. http://ataspinar.com/2018/12/21/a-guide-for-using-the-wavelet-transform-in-machine-learning.

Yuchimiuk, J. (2007). System and method for gender identification in a speech application environment. US Patent App. 10/186,049.

Zhang, K., Gao, C., Guo, L., Sun, M., Yuan, X., Han, T. X., Zhao, Z., and Li, B. (2017). Age group and gender estimation in the wild with deep ror architecture. *IEEE Access*, 5:22492–22503.

Zhenyu He (2010). Activity recognition from accelerometer signals based on wavelet-ar model. In *2010 IEEE International Conference on Progress in Informatics and Computing*, volume 1, pages 499–502.