UNIVERSITI PUTRA MALAYSIA

IMPROVED CLASS BINARIZATION MODEL WITH DATA OVERSAMPLING IN GAIT RECOGNITION

ABDUL RAFIEZ BIN ABDUL RAZIFF

FSKTM 2019 7

IMPROVED CLASS BINARIZATION MODEL WITH DATA OVERSAMPLING IN GAIT RECOGNITION

By

ABDUL RAFIEZ BIN ABDUL RAZIFF

Thesis Submitted to the School of Graduate Studies, Universiti Putra Malaysia, in Fulfilment of the Requirements for the Degree of Doctor of Philosophy

February 2019

DEDICATION

To my beloved family.

To my Father, Mother, Sister.

To the Partner, Siti Nur Animah, in this life and hereafter.

To our kids, Amzaar Rahman and Ammar Elyas

Abstract of thesis presented to the Senate of Universiti Putra Malaysia in fulfillment of the requirement for the degree of Doctor of Philosophy

IMPROVED CLASS BINARIZATION MODEL WITH DATA OVERSAMPLING IN GAIT RECOGNITION

By

ABDUL RAFIEZ BIN ABDUL RAZIFF

February 2019

Chairman : Assoc. Prof. Md Nasir Sulaiman, PhD
Faculty : Computer Science and Information Technology

Gait is a process of a complete cycle of walking that consist of two-step cycles. It can be said that gait has a high degree of biometric which means that every person has its own unique style of walking. Gait recognition using smartphone accelerometer has been widely used in many research and applications due to the cheap assembly, durability and reliability of the Inertial Measurement Unit (IMU) Microelectromechanical System (MEMS) technology. Gait recognition has been used in many areas such as biomechanics, neuro-rehabilitation, sports medicine, security and many others. Latest research achievement in gait recognition approach is the ability to sufficiently recognize a person with small variations and single data enrolment.

In the standard gait recognition, there are four main workflows or levels that include data acquisition, pre-processing, features extraction and classification. However, most of the current research is concentrated on the data acquisition and features extractions with a minimal concentration on other workflows, hence the best accuracy is not fully achieved and optimized.

In this thesis, we found several problems at the data acquisition stage, pre-processing stage, and classification stage. At the data acquisition stage, gait data is obtained from predefined places such as pocket, pouch, trousers and other parts of the body. However, due to the limitation of the clothes and culture, the mentioned places may not be suitable for smartphone placement.

At the pre-processing stage, linear interpolation is widely used by researchers in order to create a fix sampling rate between data points. However, they never examine the best interpolation rate for usage as the rate affects the number of data and this would significantly affect the overall accuracy.

At the classification stage, there are two problems that were observed. The first problem is the single classifier mapping applied by the current researchers which are not suitable because the gait recognition involved many classes and possible of overlapped classes boundary is high, hence multiclass classification or binarization of classes should be adopted. However, some researcher does apply one-vs-all (OVA) and one-vs-one (OVO) multiclass methods but the classes are not widely spread and it is not well distributed among class comparison. The second problem in the classification stage is the imbalance class when binarization dataset is performed after the multiclass classification mapping is applied.

To overcome the problems mentioned above, we proposed new methods to tackle the problems at the mentioned stages. At the data acquisition stage, we proposed a method that uses hand as the position of the smartphone. At the pre-processing stage, Linear Interpolation Factor Determinator (LIFD) is proposed by using decision tree and cross-validation evaluation in-order to determine the best linear interpolation rate. At the classification stage, we proposed the used of Random Correction Code (RCC) as the main multiclass classifier mapping. RCC is an extension of Error-correcting Output Code (ECOC) that is used for multiclass classification. To tackle the imbalance class problem, a new oversampling method, Self-adjusted Synthetic Minority Over-sampling Technique (SA-SMOTE) is proposed to automatically assign number of samples on the minority class without human intervention.

For the experimentation, gait data using hands (HHScD) is collected from 30 subjects with three different poses. Then it is investigated whether it is viable for the gait recognition process. The dataset was compared with the largest gait database from Osaka University (OU-ISIR-2) which the data was captured from smartphone clipped to the waist belt from 408 subjects. Then our proposed methods was applied to the dataset and comparison with the existing method was evaluated. Our experimental results show improvements of the accuracy in comparison with the previous study.

Abstrak tesis yang dikemukakan kepada Senat Universiti Putra Malaysia sebagai memenuhi keperluan untuk ijazah Doktor Falsafah

PENAMBAHBAIKAN KLASIFIKASI PENDUAAN DENGAN PENGSAMPELAN LEBIHAN DATA UNTUK PENGECAMAN GAIT

Oleh

ABDUL RAFIEZ BIN ABDUL RAZIFF

Februari 2019

Pengerusi : Prof. Madya Md Nasir Sulaiman, PhD
Fakulti : Sains Komputer dan Teknologi Maklumat

Gait adalah proses kitaran lengkap berjalan kaki yang terdiri daripada kitaran dua langkah. Ia boleh dikatakan bahawa *gait* mempunyai ciri biometrik yang tinggi di mana setiap orang mempunyai gaya berjalan yang unik. Pengecaman *gait* menggunakan pecutan telefon pintar telah digunakan secara meluas dalam pelbagai penyelidikan dan aplikasi kerana kos pemasangan yang murah, ketahanan dan kebolehpercayaan teknologi Mekanik Mikroelektrik (MEMS) Unit Pengukuran Inertia (IMU). Pengecaman *gait* telah digunakan dalam banyak bidang seperti biomekanik, pemulihan neuro, perubatan sukan, keselamatan dan lain-lain lagi. Pencapaian penyelidikan terkini dalam pendekatan pengiktirafan *gait* adalah keupayaan untuk mengenali seseorang dengan variasi kecil dan pendaftaran data tunggal.

Dalam pengecaman standard *gait*, terdapat empat aliran kerja utama atau tahap yang termasuk pengambilalihan data, pra-pemprosesan, ciri-ciri pengekstrakan dan klasifikasi. Walau bagaimanapun, kebanyakan penyelidikan semasa tertumpu kepada pengambilalihan data dan ciri-ciri pengekstrakan dengan kepadatan yang minimum pada alur kerja yang lain, maka ketepatan yang terbaik tidak dicapai sepenuhnya dan dioptimumkan.

Dalam tesis ini, kami mendapati beberapa masalah di peringkat pemerolehan data, tahap pra pemprosesan, dan peringkat klasifikasi. Pada peringkat pemerolehan data, data *gait* diperolehi dari tempat yang telah ditentukan seperti poket, seluar dan bahagian lain badan. Walau bagaimanapun, disebabkan oleh had pakaian dan budaya, tempat-tempat yang disebutkan itu mungkin tidak sesuai sebagai penempatan telefon pintar.

Pada tahap pra-pemprosesan, interpolasi linear digunakan secara meluas oleh para penyelidik untuk menghasilkan kadar persampelan penetapan antara titik data. Walau bagaimanapun, mereka tidak pernah memeriksa kadar interpolasi yang terbaik untuk penggunaan kerana kadarnya mempengaruhi bilangan data dan ini akan memberi kesan yang ketara kepada ketepatan keseluruhan.

Pada peringkat klasifikasi, terdapat dua masalah yang kami temui. Masalah pertama ialah pemetaan pengelasan tunggal yang digunakan oleh penyelidik semasa yang tidak sesuai kerana pengecaman *gait* melibatkan banyak kelas dan kemungkinan sempadan kelas yang bertindih adalah tinggi, oleh itu klasifikasi kelas atau kelas penduaan harus diterima pakai. Walau bagaimanapun, sesetengah penyelidik menggunakan kaedah *multiclass one-vs-all* (OVA) dan *one-vs-one* (OVO) tetapi kelas tidak tersebar luas dan tidak diagihkan dengan baik di kalangan perbandingan kelas. Masalah kedua dalam peringkat klasifikasi adalah ketidakseimbangan kelas ketika set data penduaan dilakukan setelah pemetaan pengelasan klasifikasi berbilang kelas diterapkan.

Untuk mengatasi masalah yang disebutkan di atas, kami mencadangkan kaedah baru untuk menangani masalah di peringkat yang telah disebutkan. Pada peringkat pemerolehan data, kami mencadangkan kaedah yang menggunakan tangan sebagai kedudukan telefon pintar. Pada peringkat pra-pemprosesan, Penentu Kadar Faktor Interpolasi Linear (LIFD) dicadangkan dengan menggunakan pepohon keputusan dan penilaian pengesahan silang untuk menentukan kadar interpolasi linear terbaik. Pada peringkat klasifikasi, kami mencadangkan Kod Pembetulan Rawak (RCC) yang digunakan sebagai pemetaan pengelas *multiclass* utama. RCC adalah pelanjutan Kod Output Ralat-Kesalahan (ECOC) yang digunakan untuk klasifikasi berbilang. Untuk menangani masalah ketidakseimbangan kelas, kaedah pensampelan lebihan yang baru, Teknik Penyampingan Minoriti sintetik yang diselaraskan sendiri (SA-SMOTE) dicadangkan untuk menetapkan bilangan sampel secara automatik pada kelas minoriti tanpa intervensi manusia.

Untuk eksperimen, data *gait* menggunakan tangan (HHScD) dikumpulkan dari 30 subjek dengan tiga pose yang berbeza. Kemudian ia disiasat sama ada ia layak untuk proses pengecaman *gait*. Set data ini telah dibandingkan dengan pangkalan data *gait* terbesar dari Universiti Osaka (OU-ISIR-2) yang datanya diambil dari telefon pintar diletakkan ke atas pinggang 408 subjek. Kemudian kaedah yang dicadangkan telah diterapkan pada set data dan perbandingan dengan kaedah yang sedia ada telah dinilai. Keputusan eksperimen kami menunjukkan peningkatan ketepatan berbanding dengan kajian terdahulu.

ACKNOWLEDGEMENTS

In the name of Allah, the most gracious and as the most merciful.

First and foremost, I would like to express my sincere appreciation to my supervisor, Associate Professor Dr. Md Nasir Sulaiman for giving me a tremendous support throughout my research including sponsoring me for conferences. His patience and knowledge helped me to complete this thesis.

Secondly, I would like to thank my supervisory committee: Assoc. Prof. Datin Dr. Norwati Mustapha and Dr. Thinagaran Perumal for their encouragement, precious ideas and positive comments.

I also would like to extend my appreciation to my PhD labmates especially Raihani, Shah, Nizam, Sham, Syafiq and other lab friends for their kind and moral support throughout my study in UPM.

Last but not least, I would like to express my special gratitude to my wife and sons, who are autistic. They gave me the best support, strength and encouragement throughout my study. At the same time, I also would like to thank my parents, sister and brother in law for their support.

This thesis was submitted to the Senate of Universiti Putra Malaysia and has been accepted as fulfilment of the requirement for the degree of Doctor of Philosophy. The members of the Supervisory Committee were as follows:

Md Nasir Sulaiman, PhD
Associate Professor
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Chairman)

Datin Norwati binti Mustapha, PhD
Associate Professor
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Member)

Thinagaran Perumal, PhD
Senior Lecturer
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Member)

ROBIAH BINTI YUNUS, PhD
Professor and Dean
School of Graduate Studies
Universiti Putra Malaysia

Date:

Declaration by graduate student

I hereby confirm that:

- this thesis is my original work;
- quotations, illustrations and citations have been duly referenced;
- this thesis has not been submitted previously or concurrently for any other degree at any institutions;
- intellectual property from the thesis and copyright of thesis are fully-owned by Universiti Putra Malaysia, as according to the Universiti Putra Malaysia (Research) Rules 2012;
- written permission must be obtained from supervisor and the office of Deputy Vice-Chancellor (Research and innovation) before thesis is published (in the form of written, printed or in electronic form) including books, journals, modules, proceedings, popular writings, seminar papers, manuscripts, posters, reports, lecture notes, learning modules or any other materials as stated in the Universiti Putra Malaysia (Research) Rules 2012;
- there is no plagiarism or data falsification/fabrication in the thesis, and scholarly integrity is upheld as according to the Universiti Putra Malaysia (Graduate Studies) Rules 2003 (Revision 2012-2013) and the Universiti Putra Malaysia (Research) Rules 2012. The thesis has undergone plagiarism detection software

Signature: ________________________ Date: ____________________

Name and Matric No: Abdul Rafiez bin Abdul Raziff, GS 38196

Declaration by Members of Supervisory Committee

This is to confirm that:

- the research conducted and the writing of this thesis was under our supervision;
- supervision responsibilities as stated in the Universiti Putra Malaysia (Graduate Studies) Rules 2003 (Revision 2012-2013) were adhered to.

Signature: ______________________________
Name of Chairman of Supervisory Committee: Associate Professor Dr. Md Nasir b Sulaiman

Signature: ______________________________
Name of Member of Supervisory Committee: Associate Professor Datin Dr. Norwati binti Mustapha

Signature: ______________________________
Name of Member of Supervisory Committee: Dr. Thinagaran Perumal

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Figure **Page**

LIST OF ABBREVIATIONS

CCI	Correct Classified Instances
CRM	Cycle Rotation Metric
DCT	Discrete Cosine Transform
DTW	Dynamic Time Warping
ELM	Extreame Learning Machine
FFT	Fast Fourier Transform
HHScD	Hand Held Self-collected Dataset
HOS	High Order Statistics
ICI	Incorrect Classified Instances
IMU	Inertial Measurement Unit
IRA	Identification Recognition Accuracy
k-NN	K-nearest Neighbor
LIFD	Linear Interpolation Factor Determinator
MEMS	Microelectromechanical Systems
MLP	Multilayer Perceptron Neural Network
MOS	Measure of Similarity
NCR	Number of Correct Recognition
NIR	Number of Incorrect Recognition
OU-ISIR	Osaka University and Institute of Scientific and Industrial Research
OU-ISIR-2	Osaka University and Institute of Scientific and Industrial Research Subset 2
OVA	One-vs-all
OVO	One-vs-one
PCA	Principal Component Analysis
RCC	Random Correction Code
RMS	Root Mean Square
SMOTE	Synthetic Minority Over-sampling Technique
SA-SMOTE	Self-adjust – Synthetic Minority Over-sampling Technique

SPA	Singular Spectrum Analysis
SVM	Support Vector Machine

CHAPTER 1

INTRODUCTION

1.1 Background and Motivation

Gait is a term that refers to the complete cycle of walking that involves a collection of steps. Analysis of gait is an important area as a normal human being, gait poses a unique signature of a person which can be used as a security measure like other biometrics such as fingerprints, voice, iris and face (Sun et al., 2014). It is also considered as behavioral biometric (Galbally et al., 2010). The advantage of using gait unlike other biometric are: (1) user-friendly in the data capture process which it is unobtrusive and continuous, and (2) gait is hard to be mimicked by other people (Mjaaland et al., 2010; Muaaz & Mayrhofer, 2017). In health, gait is considered as very important as it can be used to reflect human's health condition such as cerebral palsy, Parkinson's disease, cardiopathies and after effect of stroke (Azevedo Coste et al., 2014; Muro-De-La-Herran et al., 2014; Pogorelc et al., 2012).

At present, analysis of gait is done by performing gait recognition which the gait's signal data collection can be done by using two methods which are wearable and non-wearable based (Muro-De-La-Herran et al., 2014). The wearable based is the type of capturing the gait data by using a device such as an accelerometer or gyroscope. Non-wearable based is the type that does not require the user to wear anything. Usually, the data is captured from the floor sensor or by the vision based.

With the rapid development of Microelectromechanical System (MEMS), gait data can be easily collected without the needs of an expensive device. Previously, to get the gait data, one needs to use a special accelerometer device. However, nowadays smartphone does equip with many sensors and one of it is inertial measurement unit (IMU) such as accelerometer and gyroscope which was initially created for tilt movement of the device. The advantage of using this kind of device is the power consumption is low, small in size, portable and low in development cost (Deng et al., 2016; Lu et al., 2017). This is actually the goal of ubiquitous computing which is to integrate the environment, everyday objects and activities with smartphones so that it can assist people's daily life and activities (Georgievski & Aiello, 2017).

According to Statista, in Malaysia itself, the number of smartphone penetration rate as share of the population in Malaysia is 65.14% in 2018 and the number will keep on increasing to 68.46% in the year 2022 (Statista, 2018). This statistic shows that more data can be captured from the user and more analysis and pattern recognitions can be made for a lot of task including marketing, health, transportation, logistics and others.

At the current research application, there are still challenges in the field of gait recognition that needs attention in finding the best solution. The main challenges of the gait recognition can be divided into sensor induced and gait characteristics (Sprager & Juric, 2015b). Sensor induced factor is caused by the sensor itself including position and orientation. In the gait characteristics, the person or the environment do influence the gait signal such as the health of the person, physiology, clothing, walking surface and more (Gafurov et al., 2011).

Generally, in order to perform gait recognition using wearable sensor, there are stages that need to be followed which starts from the data acquisition. At this stage, the smartphone is attached to the body for the gait signal recording. Then, the next step is to perform preprocessing that includes treatments to the dataset. After that, patterns transformation can be captured by performing certain algorithms in finding descriptive information of the gait signal. After gaining informative data, recognition of gait can be performed by employing pattern similarity matching or machine learning algorithm (Sprager & Juric, 2015b).

In this thesis, few problems are found from the existing research work. The problems are found in the gait recognition stages which are: (1) data acquisition, (2) pre-processing, and (3) classification. The first problem is related to positioning of the smartphone at the data acquisition stage. In the real world application, clothes and culture may differ so positions that are used by the current researcher may not be suitable. To overcome this problem, we proposed the used of hand placement. To support our proposal, we compare with the largest gait dataset obtained from Osaka University.

At the pre-processing stage, we identified a problem related to the linear interpolation application used by current researchers. It can be seen that different sampling rate produce different number of data which subsequently affect the accuracy. To overcome this situation, we proposed a factor determinator in choosing the best sampling rate. We compare with the current sampling rate used by previous researchers to support our proposal.

At the classification stage, we observed a problem that is related to the overlapped class boundary issue due to the large number of people as class. At the same time, imbalance class problem is also noticed. To solve this problem, multiclass classification is proposed. At the same time, new oversampling method is proposed. Oversampling is a process of adding data to the existing dataset. In this work, the oversampling method is chosen instead of undersampling because when binarization is done, the goal is to retain the information on both binarized classes. Then experiment on the classification is performed and comparison with the current classification method.

The motivation of this thesis is to open new fundamentals in gait recognition especially in placing the smartphone on hand during walking for data collection. At the same time, if a person does not has any pocket or pouch as the clothes may not have it, using hand would be the best place for placement. Besides that, the interpolation rate should be analysed too before proceeding to further processes. In the classification stage, multiclass classification with automatic oversampling can be employed to increase the balance of dataset hence increasing the overall accuracy. This would save the researcher's effort especially when they need to generate more features for extraction.

1.2 Problem Statement

The first problem discusses the position of the smartphone for data collection. Various smartphone sensor placement has been used in past research such as pocket (Derawi & Bours, 2013; Hoang et al., 2015; Sun et al., 2014), pouch (Hoang et al., 2015; Trung et al., 2014; Nickel & Busch, 2013; Sun et al., 2014), clipped to the waistband of the clothes (Jordan et al., 2013), ankle (Sun & Yuao, 2012; Zhang et al., 2015), other multiple parts of body (Ren et al., 2015) and arm (Zhang et al., 2015). In the real world situation, not many people would be able to place the smartphone at the mentioned location as above due to the limitation of culture clothing especially in female clothing attire such as baju kurung, cheongsam, saree, burqa and many more. Some of ladies clothing also do not have an appropriate pocket (Melanie, 2017). Having said that, a smartphone that is handheld or on hand is never been discussed. This probe a question whether the handheld based signal is viable or not.

Another problem that arises is at the preprocessing stage which is related to the application of the linear interpolation. Since the source of the accelerometer signal is from the smartphone, the linear interpolation method is considered as one of the most favorable methods for re-sampling at a specific frequency (Derawi & Bours, 2013; Derawi et al., 2010; Gafurov et al., 2006; Hoang et al., 2015; Hoang et al., 2013; Muaaz & Mayrhofer, 2013; Muaaz & Nickel, 2012; Nickel et al., 2011; Sprager & Zazula, 2009, 2011; Thang et al., 2012). However, according to our observation, there is an issue using different frequency rate which it does affect the overall recognition accuracy as it plays the number of samples for a particular person. In the past research, there is no investigation that checks the suitability of a specific sample rate. It means that before further pre-processing or action can be taken, a suitable frequency need to be determined.

Another problem that we observed is at the classification stage which gait recognition does involve with many people which means many classes in the learning model. There are some works in gait recognition that use single classifier approach (Derawi & Bours, 2013; Hoang, et al., 2013; Nickel & Busch, 2013; Ren et al., 2015). However, using single classifier mapping does pose an overlapping class issue and most of the classifier a very well suited with the binary class problem (Dietterich & Bakiri, 1995; Fernández et al., 2010). Hence multiclass classification mapping should be used to reduce the overlapped class boundary complexity. There is a work in gait recognition using OvO multiclass classification mapping (Muaaz & Mayrhofer, 2013) but it can be seen that OvO's method do only compare with other class in binary mode instead of comparing with the combination of classes.

Another issue that arises to us is the imbalance class problem, especially when using a multiclass classification problem. In general, many researchers do apply multiclass classification mapping for classification but they do not concentrate on the imbalance class problem as when generating new pairing dataset, a number of samples for a particular class in binary dataset do tend to be low while other class may be high. This problem will affect the overall accuracy as the data distribution among classes is not well spread (Fernández et al., 2010; Jain et al., 2014).

1.3 Research Objectives

The primary objectives of this research are to propose a new binarization classification method in-order to improve the efficiency of gait recognition using handheld based signal and at the same time reducing the problems suffered in the multiclass classification. In order to achieve the objective, the following sub-objectives are adapted to form a new framework:

1. To propose a method of data acquisition that uses hands.
2. To propose a method in finding the best factor for linear interpolation.
3. To propose the application of multiclass classification to overcome the possibility of an overlapped class boundary.
4. To enhance oversampling for the binarized dataset on the current SMOTE method by adopting the automatic data requirement parameter.

1.4 Research Scope

The scope of this work is centralized on the gait identification application. A class in the dataset is referred to a person. The dataset is obtained only from the smartphone accelerometer sensor for both HHScD and OU-ISIR-2. HHScD is the dataset that captured by placing the smartphones on hand. It is divided into three types of position. OU-ISIR-2 is a dataset obtained from Osaka University that contains the largest gait database, consist of 408 subjects (Trung et al., 2014). So altogether, there will be four types of dataset that will be used to support the experimentation of the proposed method.

Besides that, this research scope is focused on the preprocessing especially at the sampling rate and classification. The experiment is compared with several previous methods in the domain area.

1.5 Research Contributions

The overall contribution of this research is to present a new framework for gait recognition and multiclass classification. The main contributions are as follows:

1. A new smartphone placement or position for gait signal data collection.

2. A new method in linear interpolation by finding the best interpolation factor before further processing.

3. Investigation on the performance OVA multiclass classification method in gait recognition

4. Proposing the application of RCC multiclass classification to overcome the possibility of overlapped classes

5. Introduction of SA-SMOTE for tackling imbalanced class issue in binarized classification method

1.6 Organization of the Thesis

The thesis is structured and organized into six chapters. The details are as follows:

Chapter 1 explains a brief introduction of the background study in gait recognition in the view of various perspectives. The current application trend and the challenges are also discussed. The problem or limitations of the current systems, objectives, the scope and contribution are explained meticulously in this chapter.

Chapter 2 discusses the related works related to the field of gait recognition. At the same time, the challenges in gait recognition is reviewed in detailed. Besides that, the preprocessing methods such as linear interpolation are also discussed. The data acquisition, gait patterns, classification methods including imbalance are also discussed and reviewed in detail.

Chapter 3 describes the overview of the conceptual research framework and the methodology used in gait recognition. In this chapter, the stages such as the proposed data collection, pre-processing, features extraction and classification are also described.

Chapter 4 presents the detailed of the proposed framework that includes new proposed smartphone placement, linear interpolation factor deteminator, multiclass classification mapping using OvO, OvA and RCC. Then, self-adjusted SMOTE method is explained in detail together with the application of the RCC multiclass classification method.

Chapter 5 explains the implementation process and all the experiments that have been conducted are highlighted. The performance analysis and validation results of the proposed method are discussed in detailed in this chapter.

Chapter 6 depicts the discussion related to the strength and limitations of the proposed methods. Conclusion and future work that can be implemented in future are also explained.

REFERENCES

Abidine, M. B., Fergani, B., & Ordonez, F. J. (2016). Effect of over-sampling versus under-sampling for SVM and LDA classifiers for activity recognition. *International Journal of Design & Nature and Ecodynamics*, *11*(3), 306–316.

Adler, A. (2003). Sample images can be independently restored from face recognition templates. In *Electrical and Computer Engineering, 2003. IEEE CCECE 2003. Canadian Conference on* (Vol. 2, pp. 1163–1166).

Aggarwal, C. C. (2015). Data *mining: the textbook*. Springer.

Ailisto, H. J., Lindholm, M., Mantyjarvi, J., Vildjiounaite, E., & Makela, S.-M. (2005). Identifying people from gait pattern with accelerometers. *Proceedings of SPIE - The International Society for Optical Engineering*, *5779*(May 2016), 7–14. https://doi.org/10.1117/12.603331

Anthony, G., Gregg, H., & Tshilidzi, M. (2007). Image classification using SVMs: one-against-one vs one-against-all. *ArXiv Preprint ArXiv:0711.2914*.

Aparna, U. R., & Paul, S. (2016). Feature selection and extraction in data mining. In *Green Engineering and Technologies (IC-GET), 2016 Online International Conference on* (pp. 1–3).

Aroquiaraj, I. L., & Thangavel, K. (2013). Mammogram image feature selection using unsupervised tolerance rough set relative reduct algorithm. In *Pattern Recognition, Informatics and Mobile Engineering (PRIME), 2013 International Conference on* (pp. 479–484).

Azevedo Coste, C., Sijobert, B., Pissard-Gibollet, R., Pasquier, M., Espiau, B., & Geny, C. (2014). Detection of freezing of gait in Parkinson disease: preliminary results. *Sensors*, *14*(4), 6819–6827.

Bächlin, M., Schumm, J., Roggen, D., & Töster, G. (2009). Quantifying gait similarity: User authentication and real-world challenge. In *International Conference on Biometrics* (pp. 1040–1049).

Bayat, A., Pomplun, M., & Tran, D. A. (2014). A study on human activity recognition using accelerometer data from smartphones. *Procedia Computer Science*, *34*, 450–457.

Bersch, S. D., Azzi, D., Khusainov, R., Achumba, I. E., & Ries, J. (2014). Sensor data acquisition and processing parameters for human activity classification. *Sensors (Basel, Switzerland)*, *14*(3), 4239–4270. https://doi.org/10.3390/s140304239

Bhargava, N., Sharma, G., Bhargava, R., & Mathuria, M. (2013). Decision tree analysis on j48 algorithm for data mining. *Proceedings of International Journal of Advanced Research in Computer Science and Software*

Engineering, *3*(6).

Bolle, R. M., Connell, J. H., Pankanti, S., Ratha, N. K., & Senior, A. W. (2013). *Guide to biometrics*. Springer Science & Business Media.

Bours, P., & Shrestha, R. (2010). Eigensteps: A giant leap for gait recognition. In *Security and Communication Networks (IWSCN), 2010 2nd International Workshop on* (pp. 1–6).

Bramer, M. (2002). Using J-pruning to reduce overfitting in classification trees. *Knowledge-Based Systems*, *15*(5), 301–308.

Bui, D. T., Ho, T. C., Revhaug, I., Pradhan, B., & Nguyen, D. B. (2014). Landslide susceptibility mapping along the national road 32 of Vietnam using GIS-based J48 decision tree classifier and its ensembles. In *Cartography from pole to pole* (pp. 303–317). Springer.

Chawla, N. V. (2009). Data mining for imbalanced datasets: An overview. In *Data mining and knowledge discovery handbook* (pp. 875–886). Springer.

Chawla, N. V, Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*, *16*, 321–357.

Chi, K.-J., & Schmitt, D. (2005). Mechanical energy and effective foot mass during impact loading of walking and running. *Journal of Biomechanics*, *38*(7), 1387–1395.

Cutting, J. E., & Kozlowski, L. T. (1977). Recognizing friends by their walk: Gait perception without familiarity cues. *Bulletin of the Psychonomic Society*, *9*(5), 353–356.

Del Pozo, G. B., Sanchez-Avila, C., De-Santos-Sierra, A., & Guerra-Casanova, J. (2012). Speed-independent gait identification for mobile devices. *International Journal of Pattern Recognition and Artificial Intelligence*, *26*(08), 1260013.

Deng, Z.-A., Wang, G., Qin, D., Na, Z., Cui, Y., & Chen, J. (2016). Continuous indoor positioning fusing WiFi, smartphone sensors and landmarks. *Sensors*, *16*(9), 1427.

Derawi, M., & Bours, P. (2013). Gait and activity recognition using commercial phones. *Computers & Security*, *39*(2), 137–144. https://doi.org/10.1016/j.cose.2013.07.004

Derawi, M. O., Bours, P., & Holien, K. (2010). Improved cycle detection for accelerometer based gait authentication. In *2010 Sixth International Conference on Intelligent Information Hiding and Multimedia Signal Processing* (pp. 312–317).

Derawi, M. O., Nickely, C., Bours, P., & Busch, C. (2010). Unobtrusive user-authentication on mobile phones using biometric gait recognition. In *Proceedings - 2010 6th International Conference on Intelligent Information Hiding and Multimedia Signal Processing, IIHMSP 2010* (pp. 306–311). https://doi.org/10.1109/IIHMSP.2010.83

Dietterich, T. G., & Bakiri, G. (1995). Solving multiclass learning problems via error-correcting output codes. *Journal of Artificial Intelligence Research*, *2*, 263–286.

Dubey, R., Zhou, J., Wang, Y., Thompson, P. M., Ye, J., Initiative, A. D. N., & others. (2014). Analysis of sampling techniques for imbalanced data: An n= 648 ADNI study. *NeuroImage*, *87*, 220–241.

Duchesnay, E., Cachia, A., Boddaert, N., Chabane, N., Mangin, J.-F., Martinot, J.-L., Zilbovicius, M. (2011). Feature selection and classification of imbalanced datasets. *NeuroImage*, *57*(3), 1003–1014. https://doi.org/10.1016/j.neuroimage.2011.05.011

Favre, J., Aissaoui, R., Jolles, B. M., de Guise, J. A., & Aminian, K. (2009). Functional calibration procedure for 3D knee joint angle description using inertial sensors. *Journal of Biomechanics*, *42*(14), 2330–2335.

Fernandez-Lopez, P., Liu-jimenez, J., Sanchez-redondo, C., & Sanchez-reillo, R. (2016). Gait Recognition Using Smartphone. *Carnahan*, 8. Retrieved from http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=7815698

Fernández, A., Del Jesus, M. J., & Herrera, F. (2010). Multi-class imbalanced data-sets with linguistic fuzzy rule based classification systems based on pairwise learning. In *Computational Intelligence for Knowledge-Based Systems Design* (pp. 89–98). Springer.

Frank, J., Mannor, S., Pineau, J., & Precup, D. (2013). Time series analysis using geometric template matching. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *35*(3), 740–754. https://doi.org/10.1109/TPAMI.2012.121

Frank, J., Mannor, S., Pineau, J., & Precup, D. (2013). Time Series Analysis Using Geometric Template Matching. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *35*(3), 740–754. https://doi.org/10.1109/TPAMI.2012.121

Frank, J., Mannor, S., & Precup, D. (2010). Activity and Gait Recognition with Time-Delay Embeddings. In *AAAI*.

Gafurov, D. (2007). A survey of biometric gait recognition: Approaches, security and challenges. *Annual Norwegian Computer Science Conference*, 19–21. https://doi.org/10.1109/TELFOR.2012.6419460

Gafurov, D., Bours, P., & Snekkenes, E. (2011). User authentication based on foot motion. *Signal, Image and Video Processing*, *5*(4), 457.

Gafurov, D., Helkala, K., & Søndrol, T. (2006). Biometric Gait Authentication Using Accelerometer Sensor. *JCP*, *1*(7), 51–59.

Gafurov, D., & Snekkenes, E. (2009). Gait recognition using wearable motion recording sensors. *EURASIP Journal on Advances in Signal Processing*, *2009*, 7.

Gafurov, D., Snekkenes, E., & Bours, P. (2007). Spoof attacks on gait authentication system. *IEEE Transactions on Information Forensics and Security*, *2*(3), 491–502.

Gafurov, D., Snekkenes, E., & Bours, P. (2010). Improved gait recognition performance using cycle matching. In *2010 IEEE 24th International Conference on Advanced Information Networking and Applications Workshops* (pp. 836–841).

Gafurov, D., Snekkenes, E., & Buvarp, T. E. (2006). Robustness of Biometric Gait Authentication Against Impersonation Attack. *On the Move to Meaningful Internet Systems 2006: OTM 2006 Workshops*, 479–488. https://doi.org/10.1007/11915034_71

Galar, M., Fernández, A., Barrenechea, E., Bustince, H., & Herrera, F. (2011). An overview of ensemble methods for binary classifiers in multi-class problems: Experimental study on one-vs-one and one-vs-all schemes. *Pattern Recognition*, *44*(8), 1761–1776.

Galbally, J., Cappelli, R., Lumini, A., Gonzalez-de-Rivera, G., Maltoni, D., Fierrez, J., Maio, D. (2010). An evaluation of direct attacks using fake fingers generated from ISO templates. *Pattern Recognition Letters*, *31*(8), 725–732.

Ganesan, K., Acharya, U. R., Chua, C. K., Lim, C. M., & Abraham, K. T. (2014). One-class classification of mammograms using trace transform functionals. *IEEE Transactions on Instrumentation and Measurement*, *63*(2), 304–311.

Garcia, S., Luengo, J., & Herrera, F. (2016a). *Data preprocessing in data mining*. Springer.

Garcia, S., Luengo, J., & Herrera, F. (2016b). García, Luengo, Herrera - 2015 - Tutorial on practical tips of the most influential data preprocessing algorithms in data mining.pdf. *Knowledge-Based Systems*. Elsevier.

George, A. S., Roy, E., Antony, A., & Job, M. (2017). An Efficient Gait Recognition System for Human Identification using Neural Networks.

Georgievski, I., & Aiello, M. (2017). Automated planning for ubiquitous computing. *ACM Computing Surveys (CSUR)*, *49*(4), 63.

Gil, D., Girela, J. L., De Juan, J., Gomez-Torres, M. J., & Johnsson, M. (2012).

Predicting seminal quality with artificial intelligence methods. *Expert Systems with Applications*, *39*(16), 12564–12573. https://doi.org/10.1016/j.eswa.2012.05.028

Håstad, J., Impagliazzo, R., Levin, L. A., & Luby, M. (1999). A pseudorandom generator from any one-way function. *SIAM Journal on Computing*, *28*(4), 1364–1396.

He, H., & Garcia, E. A. (2008). Learning from imbalanced data. *IEEE Transactions on Knowledge & Data Engineering*, (9), 1263–1284.

Hoang, T., & Choi, D. (2014). Secure and privacy enhanced gait authentication on smart phone. *The Scientific World Journal*, *2014*.

Hoang, T., Choi, D., & Nguyen, T. (2015). Gait authentication on mobile phone using biometric cryptosystem and fuzzy commitment scheme. *International Journal of Information Security*, *14*(6), 549–560.

Hoang, T., Choi, D., Vo, V., Nguyen, A., & Nguyen, T. (2013). A lightweight gait authentication on mobile phone regardless of installation error. In *IFIP International Information Security Conference* (pp. 83–101).

Hoang, T., Nguyen, T., Luong, C., Do, S., & Choi, D. (2013). Adaptive cross-device gait recognition using a mobile accelerometer. *Journal of Information Processing Systems*, *9*(2), 333–348. https://doi.org/10.3745/JIPS.2013.9.2.333

Hossin, M., & Sulaiman, M. N. (2015). A review on evaluation metrics for data classification evaluations. *International Journal of Data Mining & Knowledge Management Process*, *5*(2), 1.

Huang, G.-B., Zhu, Q.-Y., & Siew, C.-K. (2006). Extreme learning machine: theory and applications. *Neurocomputing*, *70*(1), 489–501.

Hüllermeier, E., Fürnkranz, J., Cheng, W., & Brinker, K. (2008). Label ranking by learning pairwise preferences. *Artificial Intelligence*, *172*(16), 1897–1916.

Jain, A. K., Ross, A., & Pankanti, S. (2006). Biometrics: a tool for information security. *IEEE Transactions on Information Forensics and Security*, *1*(2), 125–143.

Jain, P. A., Raut, R., & Deshmukh, P. R. (2014). Recursive Ensemble Approach for Incremental Learning of Non-Stationary Imbalanced Data. *International Journal of Computer Applications*, *98*(17).

Johnston, A. H., & Weiss, G. M. (2015). Smartwatch-based biometric gait recognition. In *Biometrics Theory, Applications and Systems (BTAS), 2015 IEEE 7th International Conference on* (pp. 1–6).

Jordan Frank, S. M. D. P., Frank, J., Mannor, S., & Precup, D. (2010). Activity and Gait Recognition with Time-Delay Embeddings. *AAAI Conference on*

Artificial Intelligence, 407–408. https://doi.org/10.1.1.224.7630

Juefei-Xu, F., Bhagavatula, C., Jaech, A., Prasad, U., & Savvides, M. (2012). Gait-id on the move: Pace independent human identification using cell phone accelerometer dynamics. In *Biometrics: Theory, Applications and Systems (BTAS), 2012 IEEE Fifth International Conference on* (pp. 8–15).

Keogh, E., Chu, S., Hart, D., & Pazzani, M. (2001). An online algorithm for segmenting time series. *Proceedings 2001 IEEE International Conference on Data Mining*, 289–296. https://doi.org/10.1109/ICDM.2001.989531

Khoshgoftaar, T. M., Gao, K., Napolitano, A., & Wald, R. (2014). A comparative study of iterative and non-iterative feature selection techniques for software defect prediction. *Information Systems Frontiers*, *16*(5), 801–822.

Kotsiantis, S., Kanellopoulos, D., Pintelas, P., & others. (2006). Handling imbalanced datasets: A review. *GESTS International Transactions on Computer Science and Engineering*, *30*(1), 25–36.

Krawczyk, B., Galar, M., Woźniak Michałand Bustince, H., & Herrera, F. (2018). Dynamic ensemble selection for multi-class classification with one-class classifiers. *Pattern Recognition*, *83*, 34–51.

Krawczyk, B., Schaefer, G., & Wozniak, M. (2013). Combining one-class classifiers for imbalanced classification of breast thermogram features. In *Computational Intelligence in Medical Imaging (CIMI), 2013 IEEE Fourth International Workshop on* (pp. 36–41).

Krawczyk, B., & Woźniak Michałand Herrera, F. (2014). Weighted one-class classification for different types of minority class examples in imbalanced data. In *Computational Intelligence and Data Mining (CIDM), 2014 IEEE Symposium on* (pp. 337–344).

Kwapisz, J. R., Weiss, G. M., & Moore, S. A. (2010). Cell phone-based biometric identification. *IEEE 4th International Conference on Biometrics: Theory, Applications and Systems, BTAS 2010*. https://doi.org/10.1109/BTAS.2010.5634532

Li, Y., Wang, X., & Qiao, F. (2011). Gait authentication based on acceleration signals of ankle. *Chinese Journal of Electronics*, *20*(3), 447–451.

Lin, W.-J., & Chen, J. J. (2012). Class-imbalanced classifiers for high-dimensional data. *Briefings in Bioinformatics*, *14*(1), 13–26.

Liu, T., Inoue, Y., & Shibata, K. (2009). Development of a wearable sensor system for quantitative gait analysis. *Measurement*, *42*(7), 978–988.

Longadge, R., & Dongre, S. (2013). Class imbalance problem in data mining review. *ArXiv Preprint ArXiv:1305.1707*.

Lu, H., Huang, J., Saha, T., & Nachman, L. (2014). Unobtrusive Gait Verification for Mobile Phones. *Proceedings of the 2014 ACM International Symposium on Wearable Computers - ISWC '14*, 91–98. https://doi.org/10.1145/2634317.2642868

Lu, Y., Wei, Y., Liu, L., Zhong, J., Sun, L., & Liu, Y. (2017). Towards unsupervised physical activity recognition using smartphone accelerometers. *Multimedia Tools and Applications*, *76*(8), 10701–10719.

Mäntyjärvi, J., Lindholm, M., Vildjiounaite, E., Mäkelä, S. M., & Ailisto, H. (2005). Identifying users of portable devices from gait pattern with accelerometers. *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, *II*, 973–976. https://doi.org/10.1109/ICASSP.2005.1415569

Melanie Radzicki Mcmanus. (2017). Why Most Women's Clothes Don't Have Real Pockets | HowStuffWorks. Retrieved March 31, 2019, from https://lifestyle.howstuffworks.com/style/fashion/trends-looks/pockets-womens-clothes.htm

Mike Elgan. (2017). Why smartwatches failed | Computerworld. Retrieved March 20, 2019, from https://www.computerworld.com/article/3219854/why-smartwatches-failed.html

Mjaaland, B. B., Bours, P., & Gligoroski, D. (2010). Walk the walk: attacking gait biometrics by imitation. In *International Conference on Information Security* (pp. 361–380).

Mohanty, A. K., Senapati, M. R., & Lenka, S. K. (2013). A novel image mining technique for classification of mammograms using hybrid feature selection. *Neural Computing and Applications*, *22*(6), 1151–1161.

Muaaz, M., & Mayrhofer, R. (2013). An Analysis of Different Approaches to Gait Recognition Using Cell Phone Based Accelerometers. *Proceedings of International Conference on Advances in Mobile Computing & Multimedia - MoMM '13*, (August), 293–300. https://doi.org/10.1145/2536853.2536895

Muaaz, M., & Mayrhofer, R. (2014). Orientation independent cell phone based gait authentication. In *Proceedings of the 12th International Conference on Advances in Mobile Computing and Multimedia* (pp. 161–164).

Muaaz, M., & Mayrhofer, R. (2017). Smartphone-based gait recognition: from authentication to imitation. *IEEE Transactions on Mobile Computing*, *16*(11), 3209–3221.

Muaaz, M., & Nickel, C. (2012). Influence of different walking speeds and surfaces on accelerometer-based biometric gait recognition. *2012 35th International Conference on Telecommunications and Signal Processing, TSP 2012 - Proceedings*, 508–512. https://doi.org/10.1109/TSP.2012.6256346

Muro-De-La-Herran, A., Garcia-Zapirain, B., & Mendez-Zorrilla, A. (2014). Gait analysis methods: An overview of wearable and non-wearable systems, highlighting clinical applications. *Sensors*, *14*(2), 3362–3394.

Nanni, L., Fantozzi, C., & Lazzarini, N. (2015). Coupling different methods for overcoming the class imbalance problem. *Neurocomputing*, *158*, 48–61.

Ngo, T. T., Makihara, Y., Nagahara, H., Mukaigawa, Y., & Yagi, Y. (2014). Orientation-compensative signal registration for owner authentication using an accelerometer. *IEICE TRANSACTIONS on Information and Systems*, *97*(3), 541–553.

Ngo, T. T., Makihara, Y., Nagahara, H., Mukaigawa, Y., & Yagi, Y. (2014). The largest inertial sensor-based gait database and performance evaluation of gait-based personal authentication. *Pattern Recognition*, *47*(1), 228–237. https://doi.org/10.1016/j.patcog.2013.06.028

Ngo, T. T., Makihara, Y., Nagahara, H., Mukaigawa, Y., & Yagi, Y. (2015). Similar gait action recognition using an inertial sensor. *Pattern Recognition*, *48*(4), 1289–1301.

Nickel, C. (2012). Accelerometer-based biometric gait recognition for authentication on smartphones.

Nickel, C., & Busch, C. (2012). Does a cycle-based segmentation improve accelerometer-based biometric gait recognition? In *Information Science, Signal Processing and their Applications (ISSPA), 2012 11th International Conference on* (pp. 746–751).

Nickel, C., & Busch, C. (2013). Classifying accelerometer data via hidden Markov models to authenticate people by the way they walk. *IEEE Aerospace and Electronic Systems Magazine*, *28*(10), 29–35. https://doi.org/10.1109/MAES.2013.6642829

Nickel, C., Busch, C., Rangarajan, S., & Mobius, M. (2011). Using Hidden Markov Models for Accelerometer-Based Biometric Gait Recognition. *Signal Processing and Its Applications CSPA 7th International Colloquium on Signal Processing and Its Applications*, 58–63. https://doi.org/10.1109/CSPA.2011.5759842

Nickel, C., Derawi, M. O., Bours, P., & Busch, C. (2011). Scenario test of accelerometer-based biometric gait recognition. In *Security and Communication Networks (IWSCN), 2011 Third International Workshop on* (pp. 15–21).

Nickel, C., Wirtl, T., & Busch, C. (2012). Authentication of smartphone users based on the way they walk using k-NN algorithm. *Proceedings of the 2012 8th International Conference on Intelligent Information Hiding and Multimedia Signal Processing, IIH-MSP 2012*, 16–20. https://doi.org/10.1109/IIH-MSP.2012.11

Park, S.-H., & Fürnkranz, J. (2007). Efficient pairwise classification. In *European Conference on Machine Learning* (pp. 658–665).

Pirker, W., & Katzenschlager, R. (2016). Gait disorders in adults and the elderly: A clinical guide. *Wiener Klinische Wochenschrift*, *129*. https://doi.org/10.1007/s00508-016-1096-4

Pogorelc, B., Bosnić, Z., & Gams, M. (2012). Automatic recognition of gait-related health problems in the elderly using machine learning. *Multimedia Tools and Applications*, *58*(2), 333–354.

Primo, A., Phoha, V. V, Kumar, R., & Serwadda, A. (2014). Context-aware active authentication using smartphone accelerometer measurements. In *Proceedings of the IEEE conference on computer vision and pattern recognition workshops* (pp. 98–105).

Quinlan, J. R. (1986). Induction of decision trees. *Machine Learning*, *1*(1), 81–106.

Ramirez-Gallego, S., Krawczyk, B., Garc\'\ia, S., & Woźniak Michałand Herrera, F. (2017). A survey on data preprocessing for data stream mining: current status and future directions. *Neurocomputing*, *239*, 39–57.

Rana, J., Arora, N., & Hiran, D. (2017). An Intensive Assessment of Factors Affecting Gait Identification and Recognition.

Ravi, N., Dandekar, N., Mysore, P., & Littman, M. L. (2005). Activity recognition from accelerometer data. In *Aaai* (Vol. 5, pp. 1541–1546).

Reeder, B., & David, A. (2016). Health at hand: a systematic review of smart watch uses for health and wellness. *Journal of Biomedical Informatics*, *63*, 269–276.

Ren, Y., Chen, Y., Chuah, M. C., & Yang, J. (2015). User Verification Leveraging Gait Recognition for Smartphone Enabled Mobile Healthcare Systems. *IEEE Transactions on Mobile Computing*, *14*(9), 1961–1974. https://doi.org/10.1109/TMC.2014.2365185

Rifkin, R., & Klautau, A. (2004). In defense of one-vs-all classification. *Journal of Machine Learning Research*, *5*(Jan), 101–141.

Rocha, A., & Goldenstein, S. K. (2014). Multiclass from binary: Expanding one-versus-all, one-versus-one and ecoc-based approaches. *IEEE Transactions on Neural Networks and Learning Systems*, *25*(2), 289–302.

Rong, L., Jianzhong, Z., Ming, L., & Xiangfeng, H. (2007). A wearable acceleration sensor system for gait recognition. In *Industrial Electronics and Applications, 2007. ICIEA 2007. 2nd IEEE Conference on* (pp. 2654–2659).

Rong, L., Zhiguo, D., Jianzhong, Z., & Ming, L. (2007). Identification of individual walking patterns using gait acceleration. In *Bioinformatics and*

Biomedical Engineering, 2007. ICBBE 2007. The 1st International Conference on (pp. 543–546).

Sadeghi, H., Allard, P., Prince, F., & Labelle, H. (2000). Symmetry and limb dominance in able-bodied gait: a review. *Gait & Posture*, *12*(1), 34–45.

Salzberg, S. L. (1994). C4. 5: Programs for machine learning by j. ross quinlan. morgan kaufmann publishers, inc., 1993. *Machine Learning*, *16*(3), 235–240.

Samà, A., Ruiz, F. J., Agell, N., Pérez-López, C., Català, A., & Cabestany, J. (2013). Gait identification by means of box approximation geometry of reconstructed attractors in latent space. *Neurocomputing*, *121*, 79–88. https://doi.org/10.1016/j.neucom.2012.12.042

Sammut, C., & Webb, G. I. (Eds.). (2017). Cross-Validation. In *Encyclopedia of Machine Learning and Data Mining* (p. 306). Boston, MA: Springer US. https://doi.org/10.1007/978-1-4899-7687-1_190

Schlachetzki, J., Eskofier, B., Winkler, J., Marxreiter, F., Henze, T., Klucken, J., Barth, J. (2013). Unbiased and Mobile Gait Analysis Detects Motor Impairment in Parkinson's Disease. *PLoS ONE*, *8*(2), e56956. https://doi.org/10.1371/journal.pone.0056956

Seel, T., Raisch, J., & Schauer, T. (2014). IMU-based joint angle measurement for gait analysis. *Sensors*, *14*(4), 6891–6909.

Sekine, M., Tamura, T., Yoshida, M., Suda, Y., Kimura, Y., Miyoshi, H., … Fujimoto, T. (2013). A gait abnormality measure based on root mean square of trunk acceleration. *Journal of Neuroengineering and Rehabilitation*, *10*(1), 118.

Shakoor, A. (2010). Soft computing based feature selection for environmental sound classifcation.

Sprager, S., & Juric, M. B. (2015a). An Efficient HOS-Based Gait Authentication of Accelerometer Data. *IEEE Transactions on Information Forensics and Security*, *10*(7), 1486–1498. https://doi.org/10.1109/TIFS.2015.2415753

Sprager, S., & Juric, M. B. (2015b). *Inertial sensor-based gait recognition: A review*. *Sensors (Switzerland)* (Vol. 15). https://doi.org/10.3390/s150922089

Sprager, S., & Zazula, D. (2009). A cumulant-based method for gait identification using accelerometer data with principal component analysis and support vector machine. *WSEAS Transactions on Signal Processing*, *5*(11), 369–378.

Sprager, S., & Zazula, D. (2011). Impact of different walking surfaces on gait identification based on higher-order statistics of accelerometer data. In *Signal and Image Processing Applications (ICSIPA), 2011 IEEE International Conference on* (pp. 360–365).

Statista. (2018). Smartphone penetration rate as share of the population in Malaysia from 2015 to 2022. Retrieved September 9, 2018, from https://www.statista.com/statistics/625418/smartphone-user-penetration-in-malaysia/

Subramanian, R., Sarkar, S., Labrador, M., Contino, K., Eggert, C., Javed, O., Cheng, H. (2015). Orientation invariant gait matching algorithm based on the Kabsch alignment. In *Identity, Security and Behavior Analysis (ISBA), 2015 IEEE International Conference on* (pp. 1–8).

Sun, B., Wang, Y., & Banda, J. (2014). Gait Characteristic Analysis and Identification Based on theiPhone's Accelerometer and Gyrometer. *Sensors*, *14*(9), 17037–17054. https://doi.org/10.3390/s140917037

Sun, H., Yuan, T., Li, X., & Hu, Y. (2012). Accelerometer-based gait authentication via neural network. *Chinese Journal of Electronics*, *21*(3), 481–484.

Sun, H., & Yuao, T. (2012). Curve aligning approach for gait authentication based on a wearable accelerometer. *Physiological Measurement*, *33*(6), 1111.

Surhone, L. M., Timpledon, M. T., & Marseken, S. F. (2010). *Pseudorandom Number Generator: Algorithm, Randomness, Monte Carlo Method, Procedural Generation, Cryptography, Linear Congruential Generator, Lagged Fibonacci Generator, Feedback with Carry Shift Registers*. Betascript Publishing. Retrieved from https://books.google.com.my/books?id=cUsXQwAACAAJ

Tahir, M. A., Kittler, J., & Yan, F. (2012). Inverse random under sampling for class imbalance problem and its application to multi-label classification. *Pattern Recognition*, *45*(10), 3738–3750.

Thang, H. M., Viet, V. Q., Dinh Thuc, N., & Choi, D. (2012). Gait identification using accelerometer on mobile phone. *2012 International Conference on Control, Automation and Information Sciences (ICCAIS)*, 344–348. https://doi.org/10.1109/ICCAIS.2012.6466615

Trivino, G., Alvarez-Alvarez, A., & Bailador, G. (2010). Application of the computational theory of perceptions to human gait pattern recognition. *Pattern Recognition*, *43*(7), 2572–2581.

Trojaniello, D., Ravaschio, A., Hausdorff, J. M., & Cereatti, A. (2015). Comparative assessment of different methods for the estimation of gait temporal parameters using a single inertial sensor: application to elderly, post-stroke, Parkinson's disease and Huntington's disease subjects. *Gait & Posture*, *42*(3), 310–316.

Trung, N. T., Makihara, Y., Nagahara, H., Mukaigawa, Y., & Yagi, Y. (2012). Performance evaluation of gait recognition using the largest inertial sensor-based gait database. In *Biometrics (ICB), 2012 5th IAPR International*

Conference on (pp. 360–366).

Trung, N. T., Makihara, Y., Nagahara, H., Sagawa, R., Mukaigawa, Y., & Yagi, Y. (2011). Phase registration in a gallery improving gait authentication. In *Biometrics (IJCB), 2011 International Joint Conference on* (pp. 1–7).

Vera-Rodriguez, R., Mason, J. S. D., Fierrez, J., & Ortega-Garcia, J. (2013). Comparative analysis and fusion of spatiotemporal information for footstep recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *35*(4), 823–834.

Wada, O., Asai, T., Hiyama, Y., Nitta, S., & Mizuno, K. (2017). Root mean square of lower trunk acceleration during walking in patients with unilateral total hip replacement. *Gait & Posture*, *58*, 19–22.

Wang, J., She, M., Nahavandi, S., & Kouzani, A. (2010). A review of vision-based gait recognition methods for human identification. In *Digital Image Computing: Techniques and Applications (DICTA), 2010 International Conference on* (pp. 320–327).

Wang, J., You, J., Li, Q., & Xu, Y. (2012). Extract minimum positive and maximum negative features for imbalanced binary classification. *Pattern Recognition*, *45*(3), 1136–1145.

Wang, L., Tan, T., Ning, H., & Hu, W. (2003). Silhouette analysis-based gait recognition for human identification. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *25*(12), 1505–1518.

Wasikowski, M., & Chen, X. (2010). Combating the small sample class imbalance problem using feature selection. *IEEE Transactions on Knowledge and Data Engineering*, *22*(10), 1388–1400.

Watanabe, Y. (2014). Influence of holding smart phone for acceleration-based gait authentication. *Proceedings - 2014 International Conference on Emerging Security Technologies, EST 2014*, 30–33. https://doi.org/10.1109/EST.2014.24

Weiss, G. M., McCarthy, K., & Zabar, B. (2007). Cost-sensitive learning vs. sampling: Which is best for handling unbalanced classes with unequal error costs? *DMIN*, *7*, 35–41.

Whittle, M. W. (2014). *Gait analysis: an introduction*. Butterworth-Heinemann.

Witten, I. H., Frank, E., Hall, M. A., & Pal, C. J. (2016). *Data Mining: Practical machine learning tools and techniques*. Morgan Kaufmann.

Youn, I.-H., Choi, S., Le May, R., Bertelsen, D., & Youn, J.-H. (2014). New gait metrics for biometric authentication using a 3-axis acceleration. In *Consumer Communications and Networking Conference (CCNC), 2014 IEEE 11th* (pp. 596–601).

Yusof, R., Kasmiran, K. A., Mustapha, A., Mustapha, N., Zin, M., & Asma, N. O. R. (2017). Techniques for Handling Imbalanced Datasets When Producing Classifier Models. *Journal of Theoretical & Applied Information Technology*, *95*(7).

Zhang, Y., Pan, G., Jia, K., Lu, M., Wang, Y., & Wu, Z. (2015). Accelerometer-based gait recognition by sparse representation of signature points with clusters. *IEEE Transactions on Cybernetics*, *45*(9), 1864–1875.

Zhao, H., Wang, Z., Qiu, S., Shen, Y., Zhang, L., Tang, K., & Fortino, G. (2018). Heading Drift Reduction for Foot-Mounted Inertial Navigation System via Multi-Sensor Fusion and Dual-Gait Analysis. *IEEE Sensors Journal*.

Zhong, Y., & Deng, Y. (2014). Sensor orientation invariant mobile gait biometrics. In *Biometrics (IJCB), 2014 IEEE International Joint Conference on* (pp. 1–8).

BIODATA OF STUDENT

Abdul Rafiez Abdul Raziff received his Bachelor of Computer Science from Universiti Putra Malaysia (UPM) and obtained his Master of Information Technology from Open University Malaysia, Malaysia (OUM). He has 13 years working experience in teaching and handling IT based projects to companies and organization. He also has 10 years experience in pattern recognition research works and tasks. With his experience and working experience, he started to pursue his study in Intelligent Computing in Universiti Putra Malaysia (UPM) 4 years back. His research interest are in biometric, sensor, activity recognition, gait recognition and data mining. His current research is on gait recognition using smartphone's accelerometer sensor.

LIST OF PUBLICATIONS

Journals

Raziff, A. R., Sulaiman, M. N., Mustapha, N., & Perumal, T. (2017). Smote and OvO Multiclass Method for Multiple Handheld Placement Gait Identification on Smartphone's Accelerometer. *Journal of Engineering and Applied Science*, *12: 374-382*. (SCOPUS)

Raziff, A. R., Mustapha, N., Perumal, T., & Sulaiman, M. N. (2017). Gait Identification using Smartphone Handheld Placement with Linear Interpolation Factor, Single Magnitude and One-vs-one Classifier Mapping. *International Journal of Intelligent Engineering & Systems*, *10(4): 70-80*. (SCOPUS)

Raziff, A. R., Sulaiman, M. N., Mustapha, N., & Perumal, T. (2017). Single Classifier, OvO, OvA and RCC Multiclass Classification Method in Handheld Based Smartphone Gait Identification. *International Conference of Applied Science and Technology*. Langkawi: UUM (SCOPUS)

Raziff, A. R., Sulaiman, M. N., Mustapha, N., & Perumal, T. (2017). Multiclass Classification Method in Handheld Based Smartphone Gait Identification. *Journal of Telecommunication, Electronic and Computer Engineering*. (SCOPUS)

Conferences

Raziff, A. R., Sulaiman, M. N., Mustapha, N., & Perumal, T. (2016). Gait Identification using One-vs-one Classifier Model. *2016 IEEE Conference of Open System (ICOS)* (pp. 71-75). Langkawi: IEEE (SCOPUS)

Raziff, A. R., Sulaiman, M. N., Mustapha, N., & Perumal, T. (2017). Single Classifier, OvO, OvA and RCC Multiclass Classification Method in Handheld Based Smartphone Gait Identification. *International Conference of Applied Science and Technology*. Langkawi: UUM (SCOPUS)

UNIVERSITI PUTRA MALAYSIA

STATUS CONFIRMATION FOR THESIS / PROJECT REPORT
AND COPYRIGHT

ACADEMIC SESSION : SECOND SEM 2018/2019

TITLE OF THESIS / PROJECT REPORT:

IMPROVED CLASS BINARIZATION MODEL WITH DATA OVERSAMPLING IN GAIT RECOGNITION

NAME OF STUDENT: ABDUL RAFIEZ BIN ABDUL RAZIFF

I acknowledge that the copyright and other intellectual property in the thesis/project report belonged to Universiti Putra Malaysia and I agree to allow this thesis/project report to be placed at the library under the following terms:

1. This thesis/project report is the property of Universiti Putra Malaysia.
2. The library of Universiti Putra Malaysia has the right to make copies for educational purposes only.
3. The library of Universiti Putra Malaysia is allowed to make copies of this thesis for academic exchange.

I declare that this thesis is classified as:

*Please tick (√)

☐	**CONFIDENTIAL**	(Contain confidential information under Official Secret Act 1972).
☐	**RESTRICTED**	(Contains restricted information as specified by the organization/institution where research was done).
☐	**OPEN ACCESS**	I agree that my thesis/project report to be published as hard copy or online open access.

This thesis is submitted for:

☐	**PATENT**	Embargo from ________(date) until _________ (date)

Approved by:

______________________ (Signature of Student)
New IC No/ Passport No.:

Date :

______________________ (Signature of Chairman of Supervisory Committee)
Name: PROF. MADYA DR. MD NASIR SULAIMAN

Date :

[Note : If the thesis is CONFIDENTIAL or RESTRICTED, please attach with the letter from the organization/institution with period and reasons for confidentially or restricted.]