

HYBRID FEATURE SELECTION TECHNIQUE FOR CLASSIFICATION OF
HUMAN ACTIVITY RECOGNITION

NORFADZLAN BIN YUSUP

A thesis submitted in fulfilment of the
requirements for the award of the degree of
Doctor of Philosophy

School of Computing
Faculty of Engineering
Universiti Teknologi Malaysia

DECEMBER 2021

DEDICATION

This thesis is lovingly dedicated to my late father, who always encouraged me to learn and seek knowledge throughout my life. *Alfatihah*. This thesis also dedicated to my mother, who has shown me so much love and support. Additionally, this thesis is a dedication to my perseverance and refusal to give up.

ACKNOWLEDGEMENT

I would like to convey my heartfelt gratitude to my main supervisor, Professor Dr. Azlan Mohd Zain, for his support, advice and beneficial criticism. Without his continual encouragement and interest, this thesis would not be what it is now. While preparing this thesis, I communicated with a variety of individuals, including scholars and academicians. They aided in the growth of my comprehension and better understanding in signal processing and machine learning. Thank you to my external examiner, Prof. Dr. Nazri Bin Mohd Nawi (UTHM) and internal examiner, Assoc.Prof. Dr. Roliana bt Ibrahim (UTM) for their constructive feedback.

Additionally, I owe gratitude to Universiti Malaysia Sarawak (UNIMAS) and Kementerian Pengajian Tinggi (KPT) for sponsoring my doctoral studies. Additionally, librarians at UTM deserve special recognition for their aid in delivering pertinent books and School of Computing and Engineering Faculty staff for their help and support.

My postgraduate friends and lab mates should also be recognised for their assistance. My heartfelt appreciation also goes out to all my colleagues and others who have assisted me on several occasions. Their perspectives and recommendations are really beneficial. Regrettably, it is impossible to mention them all in this short space. I am indebted to every member of my family for their continuous support.

ABSTRACT

Through the advancement of wearable sensors, wireless communication, and machine learning techniques, Assistive Technologies (AT) which endorse autonomous, active, and healthy lifestyles are emerging in recent years. Among these advances, Human Activity Recognition (HAR) is one of the most innovative means to support or monitor older people in their daily lives. Extensive research in the field of HAR is also essential for the enhancement of the quality of a person's health. However, misclassifications such as intra-class variation and inter-class overlap in similar activities degrade classification accuracy. To improve the recognition of daily human activities, handcrafted features of time-domain and frequency-domain are combined. However, several extracted features may not be significant in describing the activities. Therefore, this research aims to propose a hybrid feature selection technique for optimal human activities recognition. The methodology proposed for this research is the Ensemble Filter (Relief-F and mRMR) to select the most relevant and less redundant features. Although a filter feature ranking approach is commonly used in related studies, most works fail to consider the threshold limit to exclude unnecessary and redundant features. The hybrid Binary Harmony Search-Artificial Bee Colony (BHS-ABC) algorithm further evaluated the quality of the selected features in this research. Two HAR datasets using accelerometer and gyroscope sensors from the smartphone device were evaluated, covering various daily human activities. An ensemble Random Forest (RF) was used as the base classifier to evaluate the performance of the hybrid algorithm. The performance was then compared with other shallow machine learning techniques such as Support Vector Machine (SVM), k -Nearest Neighbour (k NN), and Naïve Bayes (NB). One-versus-All (OvA) class binarization technique was adopted to enhance the detection of high-class correlations by converting initial problems of multiple classifications into a set of two-class classification problems. The proposed hybrid of Ensemble Filter with BHS-ABC for this research is capable of obtaining optimum feature subsets with high accuracy. The average of SBHAR and USCHAD datasets accuracy have improved by 3.2% and 1.6%, respectively, from the benchmark results.

ABSTRAK

Melalui kemajuan sensor boleh pakai, komunikasi tanpa wayar dan teknik pembelajaran mesin, Teknologi Bantuan (AT) untuk menyokong gaya hidup yang autonomi, aktif, dan sihat berkembang pada akhir-akhir ini. Di antara kemajuan ini, Pengecaman Aktiviti Manusia (HAR) adalah salah satu cara yang paling inovatif untuk menyokong atau mengawasi warga emas dalam kehidupan seharian mereka. Penyelidikan luas dalam bidang HAR juga penting untuk peningkatan kualiti kesihatan seseorang. Walau bagaimanapun, masalah ralat klasifikasi seperti variasi antara kelas (intra-class) dan pertindihan antara kelas dalam aktiviti yang sama (inter-class) mengurangkan ketepatan klasifikasi. Untuk meningkatkan pengecaman aktiviti manusia, ciri-ciri domain masa dan domain kekerapan digabungkan. Namun begitu, beberapa ciri yang diekstrak ini mungkin gagal menggambarkan aktiviti tersebut. Tujuan penyelidikan ini adalah untuk mencadangkan teknik pemilihan ciri hibrid untuk pengecaman aktiviti manusia yang optimum. Metodologi yang dicadangkan untuk penyelidikan ini adalah Penapis Gabungan (Relief-F dan mRMR) untuk memilih ciri-ciri yang paling relevan dan kurang berulang. Walaupun pendekatan ciri penapis biasa digunakan, kebanyakan penyelidikan tidak mengambil kira nilai had ambang untuk mengecualikan ciri-ciri yang tidak perlu dan berulang. Kemudian, algoritma Binari Carian Harmoni-Koloni Lebah Buatan (BHS-ABC) digunakan untuk menilai kualiti ciri-ciri yang dipilih. Dalam penyelidikan ini, dua dataset HAR yang menggunakan sensor pecutan dan sensor giroskop dari peranti telefon pintar akan dinilai, merangkumi pelbagai aktiviti harian manusia. Hutan Rawak (RF) digunakan sebagai pengelas asas untuk menilai prestasi algoritma hibrid. Prestasi pengelas asas dibandingkan dengan teknik pembelajaran mesin lain seperti Mesin Vektor Sokongan (SVM), k-Jiran Terdekat (kNN) dan *Naïve Bayes* (NB). Teknik binarisasi kelas Satu-lawan-Semua (OvA) digunakan untuk meningkatkan pengesanan korelasi kelas tinggi dengan menukar klasifikasi pelbagai kelas menjadi klasifikasi dua kelas. Teknik pemilihan ciri hibrid Penapis Gabungan dan BHS-ABC untuk penyelidikan ini mampu memperoleh subset ciri optimum dengan ketepatan tinggi. Ketepatan purata set data SBHAR dan USCHAD masing-masing meningkat sebanyak 3.2% dan 1.6% daripada hasil penanda aras.

TABLE OF CONTENTS

	TITLE	PAGE
	DECLARATION	iii
	DEDICATION	iv
	ACKNOWLEDGEMENT	v
	ABSTRACT	vi
	ABSTRAK	vii
	TABLE OF CONTENTS	viii
	LIST OF TABLES	xii
	LIST OF FIGURES	xvi
	LIST OF ABBREVIATIONS	xviii
	LIST OF APPENDICES	xix
CHAPTER 1	INTRODUCTION	1
	1.1 Overview	1
	1.2 Problem Background	2
	1.3 Problem Statement	4
	1.4 Aims and Objectives	7
	1.5 Scopes of the Study	7
	1.6 Significance of Research	8
	1.7 Thesis organization	9
CHAPTER 2	LITERATURE REVIEW	11
	2.1 Introduction	11
	2.2 HAR Wearable Sensors	13
	2.2.1 Sensor type	14
	2.2.2 Sensor Platform	14
	2.3 Activities of Daily Living	15
	2.4 Raw Signal Data Processing	17
	2.4.1 Filtering	18

2.4.2	Segmentation Window	18
2.5	Handcraft Features for Classification	20
2.5.1	Time-domain features	20
2.5.2	Frequency-domain features	21
2.5.3	Deep learning features	21
2.6	Intra-class variations and Inter-class similarities in HAR	22
2.7	Feature Reduction and Feature Selection	28
2.7.1	Filter	28
2.7.2	Wrapper	29
2.7.3	Embedded	30
2.7.4	Hybrid	30
2.7.5	Ensemble	30
2.8	Hybrid Feature Selection	32
2.9	Classification algorithms	35
2.9.1	K-Nearest Neighbor	37
2.9.2	Naïve Bayes	39
2.9.3	Support Vector Machine (SVM)	40
2.10	Classification strategy	42
2.11	Research Gap Analysis	45
2.12	Chapter Summary	47
2.13	Summary of References	48
CHAPTER 3	RESEARCH METHODOLOGY	55
3.1	Introduction	55
3.2	Research Flow	55
3.3	Datasets	58
3.3.1	SBHAR	59
3.3.2	USC-HAD	60
3.4	Raw Signal Pre-processing	62
3.5	Dataset Pre-processing	64
3.5.1	Signal Filtering	66
3.5.2	Feature Extraction	67

3.6	Feature Segmentation and Extraction	68
3.6.1	Window Segmentation	69
3.6.2	Accelerometer and Gyroscope Feature Extraction	72
3.7	Feature Selection	74
3.8	Classification	75
3.9	Chapter Summary	78
CHAPTER 4	PROPOSED RESEARCH WORK	79
4.1	Introduction	79
4.2	Feature Selection of Proposed Technique	79
4.3	Ensemble Filter (Relief-F-mRMR)	80
4.4	Binary Harmony Search Algorithm	87
4.4.1	Initialize the Parameters for Problem and Algorithm	88
4.4.2	Initialize the Harmony Memory (HM)	88
4.4.3	Improvise a New Harmony	89
4.4.4	Update the Harmony Memory	90
4.4.5	Check the stopping criterion	90
4.5	BHS-ABC Algorithm	91
4.6	Classification	97
4.6.1	Ensemble Random Forest	98
4.6.2	One-versus-All (OvA)	101
4.6.3	One-versus-One (OvO)	103
4.6.4	Chapter Summary	105
CHAPTER 5	EXPERIMENTAL RESULTS	107
5.1	Introduction	107
5.2	Experimental Results – SBHAR Dataset	107
5.2.1	Experimental Result of Time and Frequency-Domain	109
5.2.2	Experimental Result of Filter Feature Selection	115
5.2.3	Experimental Result of Wrapper Feature Selection	122

5.2.4	Experimental Result of Ensemble Filter-BHS-ABC	132
5.3	Experimental Result on USC-HAD Dataset	140
5.3.1	Experimental Result of Time and Frequency-Domain	141
5.3.2	Experimental Results of Filter Feature Selection	145
5.3.3	Experimental Results of Wrapper Feature Selection	151
5.3.4	Experimental Results of Ensemble Filter-BHSABC	158
5.4	Chapter Summary	166
CHAPTER 6	CONCLUSION AND FUTURE WORK	167
6.1	Introduction	167
6.2	Research Findings	168
6.2.1	Proposed Time and Frequency Domain Feature Fusion	171
6.2.2	Proposed Ensemble (Relief-F-mRMR) Filter	172
6.3	Limitation of the Study	174
6.4	Recommendation for Future Works	175
	REFERENCES	179
	APPENDICES	193
	LIST OF PUBLICATIONS	205

LIST OF TABLES

TABLE NO.	TITLE	PAGE
Table 2.1	Wearable sensor HAR dataset	17
Table 2.2	Summary of strength and weaknesses of feature selection methods	49
Table 2.3	Intra-class variation and Inter-class similarity in previous HAR	50
Table 2.4	Hybrid feature selection technique	52
Table 3.1	SBHAR and USC-HAD datasets	58
Table 3.2	SBHAR list of activities	59
Table 3.3	USC-HAD list of activities	61
Table 3.4	Activities description	61
Table 3.5	Accelerometer sensors	68
Table 3.6	Gyroscope sensors	68
Table 3.7	Time-domain and Frequency-domain features	72
Table 3.8	Standard measurement for Confusion Matrix	78
Table 4.1	Harmony Search feature selection representation	87
Table 5.1	Denoted pointer of the SBHAR experimental result	108
Table 5.2	k -fold cross-validation accuracy performance	108
Table 5.3	Classification result using each feature group	110
Table 5.4	Classification result of binarization classification vs shallow classifier	111
Table 5.5	Time required in seconds	112
Table 5.6	OvA Results	113
Table 5.7	OvO Results	114
Table 5.8	Feature group results performance	114
Table 5.9	Filter feature selection experiment model parameter	115
Table 5.10	Performance of different filter subset evaluation methods	116

Table 5.11	RF-OvA strategy vs other classifier results	116
Table 5.12	Pruned Relief-F	118
Table 5.13	Accuracy results of proposed Ensemble filter	120
Table 5.14	Results of proposed Ensemble filter	121
Table 5.15	Summary of proposed Ensemble Filter	121
Table 5.16	Wrapper feature selection experiment model parameter	122
Table 5.17	GA parameter setting.	123
Table 5.18	GA feature selection results	124
Table 5.19	BPSO parameter setting	125
Table 5.20	BPSO feature selection results	126
Table 5.21	BHSA parameter setting	126
Table 5.22	BHS feature selection results	127
Table 5.23	BHSA-ABC feature selection parameter	127
Table 5.24	BHS-ABC feature selection results	128
Table 5.25	Best fitness value in metaheuristic algorithm	128
Table 5.26	Accuracy performance of wrapper technique	130
Table 5.27	Performance of metaheuristics algorithm (Precision vs Recall)	131
Table 5.28	Performance of metaheuristics algorithm (F-Measure vs MCC)	131
Table 5.29	Hybrid BHS-ABC feature selection	133
Table 5.30	Proposed Hybrid Ensemble-BHS-ABC accuracy vs other classifier	135
Table 5.31	Proposed Hybrid Ensemble Filter-BHSABC (SBHAR-FS5)	136
Table 5.32	Metaheuristics algorithm (Wrapper) experimental	137
Table 5.33	Detailed accuracy of each activities	138
Table 5.34	Benchmark results comparison (Accuracy)	139
Table 5.35	Benchmark results comparison (Precision and Recall)	139

Table 5.36	Denoted pointer of the USC-HAD experimental result	140
Table 5.37	Classification result using each feature group	141
Table 5.38	Fusion of time and frequency domain features	142
Table 5.39	Time required in seconds	143
Table 5.40	OvA Results	143
Table 5.41	OvO Results	144
Table 5.42	Feature group results performance	144
Table 5.43	Performance of different filter subset evaluation methods	145
Table 5.44	RF-OvA strategy vs other classifier results	146
Table 5.45	Pruned Relief-F	148
Table 5.46	Accuracy results of proposed Ensemble Filter	150
Table 5.47	Results of proposed Ensemble Filter	150
Table 5.48	Summary of proposed Ensemble Filter	151
Table 5.49	GA feature selection	152
Table 5.50	BPSO feature selection	152
Table 5.51	BHS feature selection	153
Table 5.52	BHS-ABC feature selection	154
Table 5.53	Best fitness value in metaheuristic algorithm	154
Table 5.54	Accuracy performance	156
Table 5.55	Performance of metaheuristics algorithm (Precision vs Recall)	157
Table 5.56	Performance of metaheuristics algorithm (F-Measure vs MCC)	157
Table 5.57	Hybrid Ensemble Filter with BHS-ABC feature selection	159
Table 5.58	Hybrid Ensemble Filter with BHS-ABC accuracy vs other classifier	161
Table 5.59	Proposed Hybrid Ensemble filter with BHS-ABC (USCHAD-FS5)	162
Table 5.60	Metaheuristics algorithm (Wrapper) experimental	163

Table 5.61	Detailed accuracy of each activities	164
Table 5.62	Benchmark results comparison (Accuracy)	165
Table 5.63	Benchmark results comparison (Precision and Recall)	166

LIST OF FIGURES

FIGURE NO.	TITLE	PAGE
Figure 1.1	Research Aim	7
Figure 2.1	Three-axis gyroscope and accelerometer sensors	13
Figure 2.2	Category of Wearable Sensors in HAR	14
Figure 2.3	Smart textile wearable sensors	15
Figure 2.4	Different types of HAR activities	16
Figure 2.5	Example of Sliding windows (SW) per frame	19
Figure 2.6	Types of machine learning	36
Figure 2.7	Supervised learning	37
Figure 2.8	Support Vector Machine	41
Figure 2.9	Hold-out strategy	42
Figure 2.10	<i>k</i> -fold cross-validation	43
Figure 2.11	Leave-One-Out strategy	44
Figure 3.1	Research Flow	57
Figure 3.2	Illustration of sensor position	58
Figure 3.3	Sensor position of SBHAR	59
Figure 3.4	Sensor position of USC-HAD	60
Figure 3.5	Raw Data of HAR dataset	63
Figure 3.6	Filtering process	63
Figure 3.7	Raw Accelerometer	64
Figure 3.8	Raw Gyroscope	65
Figure 3.9	Accelerometer and Gyroscope signals in various human activities	66
Figure 3.10	Eight seconds of raw signal	67
Figure 3.11	Eight seconds after filter applied	67
Figure 3.12	Sliding windows per frame	69
Figure 3.13	Raw Accelerometer and Gyroscope (128 readings)	70

Figure 3.14	General process of feature extraction	71
Figure 3.15	Confusion Matrix	77
Figure 4.1	Relief-F algorithm	82
Figure 4.2	mRMR algorithm	86
Figure 4.3	Flow of BHS	91
Figure 4.4	Binary Feature Selection	92
Figure 4.5	Hybrid BHS-ABC algorithm	96
Figure 4.6	Random Forest Model	99
Figure 4.7	OvA and OvO binarization strategies	101
Figure 5.1	Out-of-bag classification (OOB) error with different number of grown trees	112
Figure 5.2	Features ranking according to score	117
Figure 5.3	Accuracy on different threshold values	119
Figure 5.4	Selected features from proposed ensemble filter	120
Figure 5.5	Average fitness value of metaheuristic algorithm	129
Figure 5.6	Metaheuristic algorithm accuracy performance	130
Figure 5.7	Convergence Curve	134
Figure 5.8	Selected features from proposed hybrid Ensemble Filter with BHS-ABC	134
Figure 5.9	Accuracy performance of hybrid Ensemble Filter-BHS-ABC	136
Figure 5.10	Features ranking according to score	147
Figure 5.11	Accuracy on different threshold values	148
Figure 5.12	Selected features from proposed ensemble filter	149
Figure 5.13	Average fitness value of metaheuristic algorithm	155
Figure 5.14	Metaheuristic algorithm accuracy	156
Figure 5.15	Convergence Curve	160
Figure 5.16	Selected features from proposed hybrid ensemble filter with BHS-ABC	160
Figure 5.17	Accuracy performance of hybrid Ensemble Filter-BHS-ABC	162

LIST OF ABBREVIATIONS

ABC	-	Artificial Bee Colony
ACO	-	Ant Colony Optimization
ADL	-	Activity of Daily Living
ANN	-	Artificial Neural Network
BA	-	Bat algorithm
BHS	-	Binary Harmony Search
BPSO	-	Binary Particle Swarm Optimization
CNN	-	Convolutional Neural Network
COA	-	Cuckoo Optimization Search algorithm
COAHS	-	Cuckoo Optimization Search-Harmony Search
FPA	-	Flower Pollination algorithm
GA	-	Genetic Algorithm
GBC	-	Genetic Algorithm- Artificial Bee Colony
HAR	-	Human Activity Recognition
HSABC	-	Harmony Search-Artificial Bee Colony
IB1	-	Instance-Based 1
IMU	-	Inertial Measurement Unit
IoT	-	Internet of Things
k -NN	-	k -Nearest Neighbour
LSSU	-	Local Search Symmetrical Uncertainty
mRMR	-	Minimum Redundancy Maximum Relevancy
NB	-	Naïve Bayesian
OOB	-	Out-Of-Bag
RF	-	Random Forest
SBHAR	-	Smartphone Based Human Activity Recognition
SU	-	Symmetrical Uncertainty
SVM	-	Support Vector Machine
USC-HAD	-	University of Southern California- Human Activity Dataset

LIST OF APPENDICES

APPENDIX	TITLE	PAGE
Appendix A	Example of raw data of SBHAR dataset	193
Appendix B	Example of raw data of USC-HAD dataset	194
Appendix C	Selected Features from Proposed Ensemble Filter (SBHAR)	195
Appendix D	Selected features from Proposed Hybrid Ensemble Filter with BHS-ABC (SBHAR)	198
Appendix E	Selected Features from Proposed Ensemble Filter (USCHAD)	200
Appendix F	Selected features from Proposed Hybrid Ensemble Filter with BHS-ABC (USCHAD)	203

CHAPTER 1

INTRODUCTION

1.1 Overview

In recent years, Assistive Technology (AT) has made great strides by permitting older people or people with disability to perform tasks, which they have not been able to accomplish in the past. By definition, AT is a device or resource that facilitates the performance of tasks that are difficult or impossible for older adults or disabled persons. In older adults, such devices may be a walker to boost their balance, or amplifier to make it easier for people to listen. It might also include a magnification software for someone with a blind or reduced vision or a pushbike that helps them travel over distances that are too long to walk. In laymen's terms, AT is everything that helps humans to keep up with activities of daily living (ADLs). For some, the ability to perform simple daily tasks, such as walking and sitting, is considered essential. AT helps most people to live without long-term caring or home healthcare comfortably.

According (Chavarriaga *et al.*, 2010), there is a raising need for healthcare and AT technologies to design machine learning techniques that capable of dealing with the uncertainty prevalent in real-world implementations. Human Activity Recognition (HAR) is used mainly as an AT for older adults and healthcare in conjunction with other devices such as the Internet of Things (IoT) with the assistance of sensors, smartphones or pictures (Jobanputra *et al.*, 2019; Abdullah *et al.*, 2020; Alema Khatun and Abu Yousuf, 2020). The Internet of Things (IoT) is a growing area of technology which aims at connecting electronic monitors with vehicles, home appliances and also (mainly) people. For consumers, businesses and communities, IoT wearables are becoming increasingly popular.

Nowadays, sensors are embedded in many devices and generate signals in response to an event or activities. Sensors can be found in the physical world and in many electronic products, which is part of today's modern-day life. For example, wearable devices have sensors that generate signals in response to human activity, such as walking or running. Over the last few decades, numerous wearable sensors and devices have been developed with advances in design and computational power in signal processing. Wearable technologies are experiencing growing requirements and interests, especially in healthcare. For example, in defining health hazards, the volumes of information produced from wearable devices could be helpful.

The Human Activity Recognition (HAR) is used to enhance the quality of life of individuals in many applications. The objective of HAR is to acknowledge the user's actions and behaviours based on the setting and the sensed movement of the user. In particular, inertial HAR approaches use the Inertial Measuring Unit (IMU) to capture the motion of the user at a particular sensor location such as the wrist, ankle, or foot. With the development of sensor technology for micro-machine mechanical systems (MEMs), inertial sensors such as accelerometer, gyroscope, and magnetometers are developed and tested to perform activity recognition.

1.2 Problem Background

Human Activity Recognition (HAR) is a vast area of study that seeks to distinguish a person's specific gesture or action based on sensor data (Wang and Meng, 2018). This includes predicting a person's behaviour, which typically requires profound knowledge and techniques from signal processing to the required technical application from the raw data in order to conform to a machine learning model (Y. Wang *et al.*, 2017). The identification of human activity is considered as a challenging classification task in time series data. The misclassification of daily human activities remains one of the challenges in HAR (Wang *et al.*, 2018). A major contributing factor to this issue is due to different people behave differently for the same human daily activities such as 'walking' or 'running'. In HAR, intra-class variation happens when one activity category involves different styles of human

motion. For example, in ‘walking’ activity, a person can walk very fast or slow. Furthermore, human activity similarities exist not only in different categories (intra-class variation) but also in the same categories (inter-class similarities). Inter-class similarities happened when the activity is physically different (e.g., walking downstairs and walking upstairs) but shows similar characteristics in sensor signal forms (Bulling *et al.*, 2014). These similarities would be challenging for intelligent machines to differentiate, thus contributes to the misclassifications.

The selection of features is another significant challenge since some features are less important which may not be useful in representing the activities. Time-domain features are often used to define static activities as a more generic, but useful one such as standing, and sitting (Zubair *et al.*, 2016; Arif *et al.*, 2017) Despite that, using these features alone to define the dynamic activities (e.g., walking and running) may not be precise (Shoaib *et al.*, 2016). While the average accuracy obtained is good, an optimum number of features still have to be chosen to prevent dimensionality curses. Generally, selection strategies for features can be categorised into filter and wrappers.

The filter feature ranking approach was employed due to their less complicated nature and is capable of handling a vast number of instances (Bolón-Canedo *et al.*, 2014). Most of the research work, however, only based on selecting top- k features and did not define the threshold which discards the low-ranking feature (Doewes *et al.*, 2017; Amezzane *et al.*, 2018; Chandra, 2018; Abo El-Maaty and Wassal, 2019). Clearly, this method is not optimized since it is likely to overestimate or degrade the pruned features. There may also be some feature redundancies and a possible increase in the false detection rate.

Currently, metaheuristic-based optimization methods for solving optimization problems have been widely used in the research. A metaheuristic optimisation approaches such as Genetic Algorithm (GA), Particle Swarm Optimization (PSO), Artificial Bee Colony (ABC) and Harmony Search (HS) is used as the wrapper-based feature selection to select optimal feature subset. In comparison to the filter, wrapper methods use a classifier to maximize classification performance by choosing the

relevant features as the learning algorithm. Metaheuristic algorithms aim primarily at avoiding drawbacks from iterative learning and the local optima issue in particular. Most of the metaheuristic algorithms are designed to solve a discrete or continuous optimization problem. Feature selection is regarded as a binary optimization problem, where solutions are restricted to binary values (Too *et al.*, 2018). Therefore, the binary version of the Harmony Search Algorithm (HSA) could be employed to solve this problem. However, metaheuristic algorithms such as HSA are likely to fall into local optima problems (Alia and Mandava, 2011; Yusup *et al.*, 2019; Li *et al.*, 2020).

1.3 Problem Statement

The presence of intra-class variation and inter-class similarity among distinct activities complicates activity recognition. When behaviours that are essentially dissimilar in nature (e.g., walking up and walking down) show remarkably similar sensor signal characteristics, this is referred to as inter-class similarity. While intra-class variation refers to an activity that involves various types of motion such as walking and running. The relationship exists because walking up and walking down are fundamentally different from other stride behaviours such as walking and running. The use of deep learning such as Convolutional Neural Network (CNN) has shown exceptional accuracy and the ability to distinguish behaviours with a high degree of inter-class similarity. Regrettably, due to the sensitivity of the lesser waveform, the CNN is incapable of generating high accuracy for static activity (Rona and Cho, 2016).

Selecting the right features is complicated because some features are ineffective or even completely unnecessary for capturing the action. The time-domain features are commonly used because they are easier and more efficient for representing static behaviour. However, relying solely on these characteristics might not be sufficient to detect dynamic behaviour. Similarly, certain features may be redundant, which may increase the rate of false classification. While it achieved a reasonable level of accuracy on average, the number of features chosen is still very

high. The benefits of using feature ranking methods to pick features are that they are less complex and can accommodate a large number of instances. However, the majority of works did not specify the feature boundary at which lower ranking features are discarded. By contrast, metaheuristic optimization techniques have been widely used to solve global optimization problems. When dealing with a large number of features, the computational cost of iteration and population re-evaluation for finding an ideal parameter has increased significantly.

Therefore, the problem statement for this research is,

A hybrid feature selection technique of the Ensemble Filter and Binary Harmony Search-Artificial Bee Colony (BHS-ABC) algorithm can select optimal features and recognize human activities for better classification accuracy.

Research questions of this work are as follows:

1. *How to reliably distinguish daily human activities for better recognition?*

Time-domain features were highly useful in discriminating between static activities like standing and sitting. On the other hand, the use of frequency-domain features is extracted to enhance the identification of dynamic activities. Time-domain features or statistical features are usually employed to distinguish the human activity especially for static or postural transition activities. However, if the frequency-domain features only are used, it could not produce high recognition performance, especially when distinguishing the static activity. Both feature categories (time and frequency domain) are, therefore, useful to distinguish between static and dynamic activities.

2. *How to select the most relevant and less redundant features based on the filter feature selection method efficiently?*

In filter feature selection, top- k features are usually employed to select the most top features rank (e.g., top 10, 50 features) based on the calculated and ranked feature weight. However, this practice is not an optimized technique since it could overestimate or underestimate the pruned features. Relief-F is an efficient feature ranking score; however, it cannot detect redundant features, which limits the application. Another filtering technique, Maximum Relevance Minimum Redundancy (mRMR), will then be used to discard these redundant features. mRMR attempts to pick a subset of features that correlate the most with a class (relevance) and the least with each other (redundancy). The selected feature subset based on the feature weight calculated from Relief-F will be the feature space for the mRMR algorithm. Then, a greedy search will be used by mRMR in order to find near-optimal features.

3. *How to find the optimal feature subset based on Binary Harmony Search-Artificial Bee Colony algorithm?*

While the BHS algorithm does not work in numerical optimisation in local search, the algorithm performance is good to identify the high-performance region of a solution space in a reasonable time. Because the ABC algorithm performs better in local and global optimisation, the binary HS algorithm hybridises by using the ABC algorithm to optimise the HS algorithm's Harmony Memory (HM). Harmony memory (HM) is a collection of the best solution and play a key role in the algorithm. ABC will be applied to optimize harmony memory as a learning mechanism. The harmony memory has considered food sources and is explored and exploited by the employed bees, the onlooker bees, and the scout bees. Then, a new harmony vector is generated in finding better solutions. If the newly improvised harmony is better, it will be stored in HM.

1.4 Aims and Objectives

This research aims to develop a hybrid filter-wrapper feature selection method using Ensemble Filter with BHS-ABC algorithm in order to get high classification accuracy with a minimal number of selected features, as shown in Figure 1.1

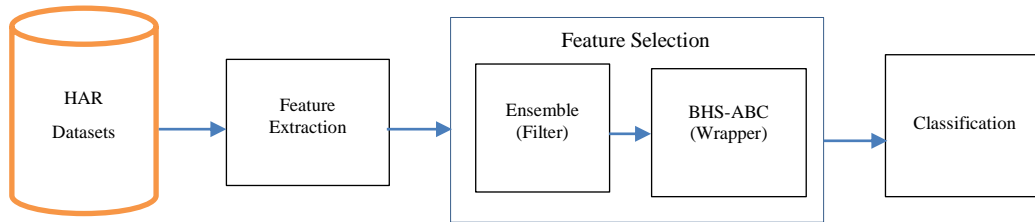


Figure 1.1 Research Aim

The objectives of this research are:

1. To fuse handcrafted time-domain and frequency-domain features for better recognition of daily human activities.
2. To develop an Ensemble Filter (Relief-F and mRMR) method to guide the search for the most relevant and less redundant feature subsets.
3. To develop a hybridization strategy by combining Ensemble Filter and Binary Harmony Search-Artificial Bee Colony algorithm for improving
4. recognition of intra-class variation and inter-class similarities activities.

1.5 Scopes of the Study

This research concentrates mainly on improving the identification of human activity behaviour, particularly when distinguishing related human activities. Two freely accessible physical activities datasets; SBHAR and USC-HAD been used in this research. These two datasets were chosen because they offer information about fundamental human behaviours, including static and dynamic activities. Additionally, the datasets were chosen because it utilised two sensors (accelerometer

and gyroscope) within the Android smartphone and Inertial Measurement Unit (IMU), both of which are attached to a single position of the human's body. This thesis does not answer the issues under environmental conditions in real time.

In both data sets, different types of static and dynamic activities are included without focusing on the transition of two or more actions. The details of static and dynamic activities used in this research are available in Chapter 3. For each dataset the experiment is performed independently, and each experiment is compared with the experimental configurations of the selected research in the literature. The proposed feature selection is evaluated separately for each data set. The benchmark works from the literature are chosen and compared with the experimental result.

1.6 Significance of Research

This study is carried out to classify numerous human behaviours based on the reported data stream from an accelerometer and gyroscope sensors, as previously mentioned. This research primarily leads to an improved identification of intra-class variation and inter-class similarity of human behaviours by using a minimum number of features. The findings of this study contribute to the theoretical process of human activity recognition at the feature extraction, feature selection, and classification stages. In the feature extraction stage, a fusion of time-domain features and frequency domain features were utilized for a better recognition of daily human activities. In feature selection stage, a new Ensemble Filter (Relief-F and mRMR) method was developed where a threshold value was introduced in Relief-F to identify the optimal amount of feature subsets. Later, an ensemble filter was developed to eliminate redundant features using the mRMR algorithm. To improve the learning process of HM in BHSA, BHSA-ABC algorithm was developed. A new two-stage hybridization technique by combining Ensemble Filter with Binary HSA-ABC was developed to improve recognition of intra-class variation and inter-class similarities activities. Class binarization strategies of One versus All (OvA) with Random Forest (RF) algorithm was also employed in the classification stage. To help

reduce the number of trees in the RF method, a self-adjusted tree was also implemented.

Human activity recognition has the potential to provide significant benefits to society, particularly in real-world such as aged and medical care. In the real world, smartphone-based evaluations have been proposed as a means of complacently monitoring gait and mobility in patients with early-stage Parkinson's disease. Patient or elderly people could benefit from a HAR system that monitors activity patterns and acts if something unusual happens, such as a change in behaviour. A HAR system may be able to help these people live more independently in the future.

1.7 Thesis organization

The thesis is made up of six chapters, each with a clear focus.

Chapter 1 provides a concise overview of activity recognition from a number of views. Furthermore, the latest trends and problems in activity identification have been analysed from numerous angles. The breadth and contribution of this research are described in this chapter, along with its limitations and research aims.

Chapter 2 details the prior work in the area of activity recognition, with specific information about the type of sensor utilised, as well as a description of earlier work using wearable sensors. Additionally, the approach of feature extraction and selection that is typically used to solve classification problems is presented.

The conceptual research framework approach for proposed better activity recognition is explained in Chapter 3. In the course of doing the activity recognition, some compulsory phases are implemented, including the pre-processing stage, the feature extraction stage, the feature selection stage, and the classification stage.

Chapter 4 discusses in depth the suggested feature selection approaches utilising Ensemble Filters (Relief-f and mRMR) with the BHS-ABC algorithm. The classification algorithms, which use OvA and OvO, have been developed to help separate activities with considerable intra-class variance from inter-class similarities.

The process and the results of the data's testing and analysis are outlined in Chapter 5. All experiments that are conducted in order to achieve the highest possible accuracy performance are analysed. The criteria for selection and the optimal parameter setting are discussed. The report also compares outcomes to previously published study.

Chapter 6 concludes the research by showing that the issue revealed in the research has been solved and is consistent with the stated objective. The report also includes a recommendation for future research based on the findings.

REFERENCES

- Abdullah, C. S., Kawser, M., Islam Opu, M. T., Faruk, T. and Islam, M. K. (2020) ‘Human Fall Detection using Built-in Smartphone Accelerometer’, in *Proceedings of 2020 IEEE International Women in Engineering (WIE) Conference on Electrical and Computer Engineering, WIECON-ECE 2020*, pp. 372–375.
- Abo El-Maaty, A. M. and Wassal, A. G. (2019) ‘Hybrid GA-PCA Feature Selection Approach for Inertial Human Activity Recognition’, *Proceedings of the 2018 IEEE Symposium Series on Computational Intelligence, SSCI 2018*. IEEE, pp. 1027–1032.
- Afkhami, S., R. Marouzi, O. and Soleimani, A. (2013) ‘A Binary Harmony Search Algorithm for Solving the Maximum Clique Problem’, *International Journal of Computer Applications*, 69(12), pp. 38–43.
- Akhavian, R. and Behzadan, A. H. (2015) ‘Construction equipment activity recognition for simulation input modeling using mobile sensors and machine learning classifiers’, *Advanced Engineering Informatics*. Elsevier Ltd, 29(4), pp. 867–877.
- Alema Khatun, M. and Abu Yousuf, M. (2020) ‘Human Activity Recognition Using Smartphone Sensor Based on Selective Classifiers’, in *2020 2nd International Conference on Sustainable Technologies for Industry 4.0, STI 2020*.
- Alia, O. M. D. and Mandava, R. (2011) ‘The variants of the harmony search algorithm: An overview’, *Artificial Intelligence Review*, 36(1), pp. 49–68.
- Alomari, O. A., Khader, A. T., Al-Betar, M. A., Abdi, Z. and Alyasseri, A. (2018) ‘A hybrid filter-wrapper gene selection method for cancer classification *’, *2018 2nd International Conference on BioSignal Analysis, Processing and Systems (ICBAPS)*. IEEE, pp. 113–118.
- Alshamlan, H. M., Badr, G. H. and Alohal, Y. A. (2015) ‘Genetic Bee Colony (GBC) algorithm: A new gene selection method for microarray cancer classification’, *Computational Biology and Chemistry*. Elsevier, 56, pp. 49–60.

- Alzubi, R., Ramzan, N., Alzoubi, H. and Amira, A. (2017) ‘A Hybrid Feature Selection Method for Complex Diseases SNPs’, *IEEE Access*, 6, pp. 1292–1301.
- Amezzane, I., Fakhri, Y., El Aroussi, M. and Bakhouya, M. (2018) ‘Analysis and Effect of Feature Selection Over Smartphone-Based Dataset for Human Activity Recognition’, *Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering, LNICST*, 206, pp. 214–219.
- Anguita, D., Ghio, A., Oneto, L., Parra, X. and Reyes-Ortiz, J. L. (2013) ‘A public domain dataset for human activity recognition using smartphones’, *ESANN 2013 proceedings, 21st European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning*, (April), pp. 437–442.
- Apolloni, J., Leguizamón, G. and Alba, E. (2016) ‘Two hybrid wrapper-filter feature selection algorithms applied to high-dimensional microarray experiments’, *Applied Soft Computing Journal*. Elsevier B.V., 38, pp. 922–932.
- Arif, M., Bilal, M., Kattan, A. and Ahamed, S. I. (2014) ‘Better physical activity classification using smartphone acceleration sensor’, *Journal of Medical Systems*, 38(9).
- Arif, M. and Kattan, A. (2015) ‘Physical activities monitoring using wearable acceleration sensors attached to the body’, *PLoS ONE*, 10(7), pp. 1–16.
- Arif, M., Kattan, A. and Ahamed, Sheikh I. (2017) ‘Classification of physical activities using wearable sensors’, *Intelligent Automation and Soft Computing*, 23(1), pp. 21–30.
- Arif, M., Kattan, A. and Ahamed, Sheikh I (2017) ‘Classification of physical activities using wearable sensors’, *Intelligent Automation and Soft Computing*, 23(1), pp. 21–30.
- Arshad, A., Zheng, G., Kim, C.-G. and Kim, S.-D. (2014) ‘Human Activity Recognition in WSN: A Comparative Study’, *International Journal of Network and Distributed Computing*, 02(04), pp. 221–230.
- Arshad, S., Feng, C., Liu, Y., Hu, Y., Yu, R., Zhou, S. and Li, H. (2017) ‘Wi-chase: A WiFi based human activity recognition system for sensorless environments’, *18th IEEE International Symposium on A World of Wireless, Mobile and Multimedia Networks, WoWMoM 2017 - Conference*. IEEE, pp. 2–7.

- Arunkumar, C. and Ramakrishnan, S. (2016) ‘Hybrid Feature Selection Using Correlation Coefficient and Particle Swarm Optimization on Microarray Gene Expression Data’, *Advances in Intelligent Systems and Computing*, 424.
- Attal, F., Mohammed, S., Dedabrishvili, M., Chamroukhi, F., Oukhellou, L. and Amirat, Y. (2015) ‘Physical human activity recognition using wearable sensors’, *Sensors (Switzerland)*, 15(12), pp. 31314–31338.
- Aziz, R., Verma, C. K. and Srivastava, N. (2017) ‘A novel approach for dimension reduction of microarray’, *Computational Biology and Chemistry*. Elsevier Ltd, 71(November), pp. 161–169.
- Bao, L. and Intille, S. S. (2004) ‘Activity recognition from user-annotated acceleration data’, *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 3001, pp. 1–17.
- Bayat, A., Pomplun, M. and Tran, D. A. (2014) ‘A study on human activity recognition using accelerometer data from smartphones’, *Procedia Computer Science*. Elsevier Masson SAS, 34, pp. 450–457.
- Belouch, M., Elhadaj, S. and Idhammad, M. (2018) ‘A hybrid filter-wrapper feature selection method for DDoS detection in cloud computing’, *Intelligent Data Analysis*, 22(6), pp. 1209–1226.
- Bolón-Canedo, V. and Alonso-Betanzos, A. (2019) ‘Ensembles for feature selection: A review and future trends’, *Information Fusion*. Elsevier, 52, pp. 1–12.
- Bolón-Canedo, V., Sánchez-Marroño, N. and Alonso-Betanzos, A. (2014) ‘Data classification using an ensemble of filters’, *Neurocomputing*, 135, pp. 13–20.
- Boser, B. E., Guyon, I. M. and Vapnik, V. N. (1992) ‘A Training Algorithm for Optimal Margin Classifiers’, in *Proceedings of the Fifth Annual Workshop on Computational Learning Theory*. New York, NY, USA: ACM (COLT '92), pp. 144–152.
- Bostani, H. and Sheikhan, M. (2017) ‘Hybrid of binary gravitational search algorithm and mutual information for feature selection in intrusion detection systems’, *Soft Computing*. Springer Berlin Heidelberg, 21(9), pp. 2307–2324.
- Bulling, A., Blanke, U. L. F. and Schiele, B. (2014) ‘A Tutorial on Human Activity Recognition Using Body-Worn’, 46(3), pp. 1–33.
- Catal, C., Tufekci, S., Pirmitt, E. and Kocabag, G. (2015) ‘On the use of ensemble of classifiers for accelerometer-based activity recognition’, *Applied Soft*

- Computing Journal*. Elsevier B.V., 37, pp. 1018–1022.
- Chandra, B. (2018) ‘An efficient feature selection technique for gene expression data’, *2018 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology, CIBCB 2018*. IEEE, pp. 1–6.
- Chaudhry, M. U. and Lee, J.-H. (2018) ‘Feature Selection For High Dimensional Data Using Monte Carlo Tree Search’, *IEEE Access*. IEEE, 6, pp. 1–1.
- Chavarriaga, R., Sagha, H., Bayati, H., Millán, J. del R., Roggen, D., Förster, K., Calatroni, A., Tröster, G., Lukowicz, P., Bannach, D., Kurz, M., Hölzl, G. and Ferscha, A. (2010) ‘Robust activity recognition for assistive technologies: Benchmarking machine learning techniques’, *Workshop on Machine Learning for Assistive Technologies at the Twenty-fourth Annual Conference on Neural Information Processing Systems (NIPS)*, (December), pp. 1–8.
- Chavarriaga, R., Sagha, H., Calatroni, A., Digumarti, S. T., Tröster, G., Millán, J. D. R. and Roggen, D. (2013) ‘The Opportunity challenge: A benchmark database for on-body sensor-based activity recognition’, *Pattern Recognition Letters*, 34(15), pp. 2033–2042.
- Chen, C., Jafari, R. and Kehtarnavaz, N. (2015) ‘UTD-MHAD: A multimodal dataset for human action recognition utilizing a depth camera and a wearable inertial sensor’, *Proceedings - International Conference on Image Processing, ICIP*. IEEE, 2015-Decem, pp. 168–172.
- Chen, G., Zhang, D., Zhu, W., Tao, Q., Zhang, C. and Ruan, J. (2012) ‘On Optimal Feature Selection Using Harmony Search for Image Steganalysis’, *2012 8th International Conference on Natural Computation, (Icnc)*, pp. 1074–1078.
- Chereshnev, R. and Kertész-Farkas, A. (2018) ‘HuGaDB: Human gait database for activity recognition from wearable inertial sensor networks’, *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 10716 LNCS, pp. 131–141.
- Chhikara, R. R., Sharma, P. and Singh, L. (2016) ‘A hybrid feature selection approach based on improved PSO and filter approaches for image steganalysis’, *International Journal of Machine Learning and Cybernetics*. Springer Berlin Heidelberg, 7(6), pp. 1195–1206.
- Chuang, L., Ke, C. and Yang, C. (2008) ‘Chuang Etal 2008’, pp. 146–150.

- Cortes, C. and Vapnik, V. (1995) ‘Support-Vector Networks’, *Machine Learning*, 20(3), pp. 273–297.
- Das, S., Singh, P. K., Bhowmik, S., Sarkar, R. and Nasipuri, M. (2016) ‘A Harmony Search Based Wrapper Feature Selection Method for Holistic Bangla Word Recognition’, *Procedia Computer Science*. The Author(s), 89, pp. 395–403.
- Davila, J. C., Cretu, A. M. and Zaremba, M. (2017) ‘Wearable sensor data classification for human activity recognition based on an iterative learning framework’, *Sensors (Switzerland)*, 17(6), pp. 1–18.
- Doewes, A., Swasono, S. E. and Harjito, B. (2017) ‘Feature selection on Human Activity Recognition dataset using Minimum Redundancy Maximum Relevance’, *2017 IEEE International Conference on Consumer Electronics - Taiwan, ICCE-TW 2017*. IEEE, (1), pp. 171–172.
- Duarte, F., Lourenço, A. and Abrantes, A. (2014) ‘Classification of Physical Activities Using a Smartphone: Evaluation Study Using Multiple Users’, *Procedia Technology*, 17, pp. 239–247.
- Fan, S., Jia, Y. and Liu, J. (2019) ‘Feature Selection of Human Activity Recognition Based on Tri-axial Accelerometer’, *Yingyong Kexue Xuebao/Journal of Applied Sciences*, 37(3), pp. 427–436.
- Feng, Z., Mo, L. and Li, M. (2015) ‘A Random Forest-based ensemble method for activity recognition’, *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*. IEEE, 2015-Novem, pp. 5074–5077.
- Guiry, J. J., van de Ven, P., Nelson, J., Warmerdam, L. and Riper, H. (2014) ‘Activity recognition with smartphone support’, *Medical Engineering and Physics*. Institute of Physics and Engineering in Medicine, 36(6), pp. 670–675.
- Guo, J., White, J., Wang, G., Li, J. and Wang, Y. (2011) ‘A genetic algorithm for optimized feature selection with resource constraints in software product lines’, *Journal of Systems and Software*. Elsevier Inc., 84(12), pp. 2208–2221.
- Hammami, M., Bechikh, S., Hung, C. C. and Ben Said, L. (2018) ‘A Multi-objective hybrid filter-wrapper evolutionary approach for feature selection’, *Memetic Computing*. Springer Berlin Heidelberg.
- Hancer, E. (2018) ‘Differential evolution for feature selection: a fuzzy wrapper–filter

- approach’, *Soft Computing*. Springer Berlin Heidelberg, (Jolliffe 2014).
- Harb, H. M. and Desuky, A. S. (2014) ‘Feature Selection on Classification of Medical Datasets based on Particle Swarm Optimization’, *International Journal of Computer Applications*, 104(5), pp. 975–8887.
- Hassan, M. M., Uddin, M. Z., Mohamed, A. and Almogren, A. (2018) ‘A robust human activity recognition system using smartphone sensors and deep learning’, *Future Generation Computer Systems*. Elsevier B.V., 81, pp. 307–313.
- Hu, Z., Bao, Y., Xiong, T. and Chiong, R. (2015) ‘Hybrid filter–wrapper feature selection for short-term load forecasting’, *Engineering Applications of Artificial Intelligence*. Pergamon, 40, pp. 17–27.
- Huda, S., Yearwood, J., Jelinek, H. F., Hassan, M. M., Fortino, G. and Buckland, M. (2016) ‘A Hybrid Feature Selection with Ensemble Classification for Imbalanced Healthcare Data: A Case Study for Brain Tumor Diagnosis’, *IEEE Access*. IEEE, 4, pp. 9145–9154.
- Inbarani, H. H., Bagyamathi, M. and Azar, A. T. (2015) ‘A novel hybrid feature selection method based on rough set and improved harmony search’, *Neural Computing and Applications*. Springer London, 26(8), pp. 1859–1880.
- Jiang, Y., Liu, X., Yan, G. and Xiao, J. (2018) ‘Modified Binary Cuckoo Search for Feature Selection: A Hybrid Filter-Wrapper Approach’, *Proceedings - 13th International Conference on Computational Intelligence and Security, CIS 2017*, 2018-Janua(2), pp. 488–491.
- Jobanputra, C., Bavishi, J. and Doshi, N. (2019) ‘Human activity recognition: A survey’, *Procedia Computer Science*. Elsevier B.V., 155(2018), pp. 698–703.
- Khan, A. M. (2011) ‘Human Activity Recognition Using A Single Tri-axial Accelerometer’, *Computer Engineering*, 93, pp. 1379–1383.
- Khan, P., Reddy, B. S. K., Pandey, A., Kumar, S. and Youssef, M. (2020) ‘Differential Channel State Information based Human Activity Recognition in IoT Networks’, *IEEE Internet of Things Journal*, 7(11), pp. 1–1.
- Khusainov, R., Azzi, D., Achumba, I. E. and Bersch, S. D. (2013) ‘Real-time human ambulation, activity, and physiological monitoring: Taxonomy of issues, techniques, applications, challenges and limitations’, *Sensors (Switzerland)*, 13(10), pp. 12852–12902.
- Kononenko, I. (1994) ‘Estimating attributes: Analysis and extensions of RELIEF’, in

- Bergadano, F. and De Raedt, L. (eds) *Machine Learning: ECML-94*. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 171–182.
- Kwapisz, J. R., Weiss, G. M. and Moore, S. A. (2011) ‘Activity recognition using cell phone accelerometers’, *ACM SIGKDD Explorations Newsletter*, 12(2), pp. 74–82.
- Van Landeghem, S., Abeel, T., Saeys, Y. and Van De Peer, Y. (2011) ‘Discriminative and informative features for biomolecular text mining with ensemble feature selection’, *Bioinformatics*, 27(13), pp. i554–i560.
- Leutheuser, H., Schuldhaus, D. and Eskofier, B. M. (2013) ‘Hierarchical, Multi-Sensor Based Classification of Daily Life Activities: Comparison with State-of-the-Art Algorithms Using a Benchmark Dataset’, *PLoS ONE*, 8(10).
- Li, S., Man, Z. and Chen, Y. (2020) ‘A self-adaptive global harmony search based extreme learning machine for classification problem’, in *International Conference on Advanced Mechatronic Systems, ICAMechS*, pp. 186–191.
- Lima, W. S., Souto, E., El-Khatib, K., Jalali, R. and Gama, J. (2019) ‘Human activity recognition using inertial sensors in a smartphone: An overview’, *Sensors (Switzerland)*, 19(14), pp. 14–16.
- Liu, H., Liu, L. and Zhang, H. (2010) ‘Ensemble gene selection by grouping for microarray data classification’, *Journal of Biomedical Informatics*. Elsevier Inc., 43(1), pp. 81–87.
- Lu, H., Chen, J., Yan, K., Jin, Q., Xue, Y. and Gao, Z. (2017) ‘A hybrid feature selection algorithm for gene expression data classification’, *Neurocomputing*. Elsevier B.V., 256(2017), pp. 56–62.
- Lumini, A., Nanni, L. and Brahnam, S. (2017) ‘Ensemble of texture descriptors and classifiers for face recognition’, *Applied Computing and Informatics*. King Saud University, 13(1), pp. 79–91.
- Mandal, I., Happy, S. L., Behera, D. P. and Routray, A. (2014) ‘A framework for human activity recognition based on accelerometer data’, *Proceedings of the 5th International Conference on Confluence 2014: The Next Generation Information Technology Summit*. IEEE, pp. 600–603.
- Masum, A. K. M., Bahadur, E. H., Shan-A-Alahi, A., Uz Zaman Chowdhury, M. A., Uddin, M. R. and Al Noman, A. (2019) ‘Human Activity Recognition Using Accelerometer, Gyroscope and Magnetometer Sensors: Deep Neural Network Approaches’, *2019 10th International Conference on Computing*,

- Communication and Networking Technologies, ICCCNT 2019*. IEEE, pp. 6–11.
- Mathews, S. M., Kambhamettu, C. and Barner, K. E. (2017) ‘Centralized class specific dictionary learning for wearable sensors based physical activity recognition’, *2017 51st Annual Conference on Information Sciences and Systems, CISS 2017*. IEEE.
- Micucci, D., Mobilio, M. and Napolitano, P. (2017) ‘UniMiB SHAR: A dataset for human activity recognition using acceleration data from smartphones’, *Applied Sciences (Switzerland)*, 7(10).
- Mohino-Herranz, I., Gil-Pita, R., Rosa-Zurera, M. and Seoane, F. (2019) ‘Activity recognition using wearable physiological measurements: Selection of features from a comprehensive literature study’, *Sensors (Switzerland)*, 19(24).
- Nahar, J., Ali, S. and Chen, Y.-P. P. (2007) ‘Microarray data classification using automatic SVM kernel selection.’, *DNA and cell biology*, 26(10), pp. 707–12.
- Nakariyakul, S. (2018) ‘High-dimensional hybrid feature selection using interaction information-guided search’, *Knowledge-Based Systems*. Elsevier B.V., 145, pp. 1–14.
- Niazi, A. H., Yazdansepas, D., Gay, J. L., Maier, F. W., Ramaswamy, L., Rasheed, K. and Buman, M. P. (2017) ‘Statistical analysis of window sizes and sampling rates in human activity recognition’, *HEALTHINF 2017 - 10th International Conference on Health Informatics, Proceedings; Part of 10th International Joint Conference on Biomedical Engineering Systems and Technologies, BIOSTEC 2017*, 5(Biostec), pp. 319–325.
- Oshiro, T. M., Perez, P. S. and Baranauskas, J. A. (2012) ‘How Many Trees in a Random Forest?’, in Perner, P. (ed.) *Machine Learning and Data Mining in Pattern Recognition*. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 154–168.
- Pati, S. K. and Das, A. K. (2013) ‘Gene Selection and Classification Rule Generation for Microarray Dataset’, pp. 73–83.
- Peng, H., Long, F. and Ding, C. (2005) ‘Feature selection based on mutual information: Criteria of Max-Dependency, Max-Relevance, and Min-Redundancy’, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(8), pp. 1226–1238.
- Pirooznia, M. and Deng, Y. (2006) ‘SVM Classifier - A comprehensive java

- interface for support vector machine classification of microarray data', *BMC Bioinformatics*, 7(SUPPL.4).
- Pirooznia, M., Yang, J. Y., Yang, M. Q. and Deng, Y. (2008) 'A comparative study of different machine learning methods on microarray gene expression data', *BMC Genomics*, 9(Suppl 1), p. S13.
- Ponce, H., Martínez-Villaseñor, M. D. L. and Miralles-Pechuán, L. (2016) 'A novel wearable sensor-based human activity recognition approach using artificial hydrocarbon networks', *Sensors (Switzerland)*, 16(7).
- Putra, I. P. E. S. and Vesilo, R. (2018) 'Window-size impact on detection rate of wearable-sensor-based fall detection using supervised machine learning', *2017 IEEE Life Sciences Conference, LSC 2017*, 2018-Janua, pp. 21–26.
- Radovic, M., Ghalwash, M., Filipovic, N. and Obradovic, Z. (2017) 'Minimum redundancy maximum relevance feature selection approach for temporal gene expression data', *BMC Bioinformatics*. *BMC Bioinformatics*, 18(1), pp. 1–14.
- Ramos, C. C. O., Souza, A. N., Chiachia, G., Falcão, A. X. and Papa, J. P. (2011) 'A novel algorithm for feature selection using Harmony Search and its application for non-technical losses detection', *Computers and Electrical Engineering*.
- Reiss, A. and Stricker, D. (2012) 'Introducing a new benchmarked dataset for activity monitoring', *Proceedings - International Symposium on Wearable Computers, ISWC*. IEEE, pp. 108–109.
- Reyes-Ortiz, J. L., Oneto, L., Samà, A., Parra, X. and Anguita, D. (2016) 'Transition-Aware Human Activity Recognition Using Smartphones', *Neurocomputing*, 171, pp. 754–767.
- Ronao, C. A. and Cho, S. B. (2016) 'Human activity recognition with smartphone sensors using deep learning neural networks', *Expert Systems with Applications*. Elsevier Ltd, 59, pp. 235–244.
- Ryu, S.-J. and Kim, J.-H. (2014) 'An Evolutionary Feature Selection Algorithm for Classification of Human Activities', in Kim, J.-H., Matson, E. T., Myung, H., Xu, P., and Karray, F. (eds) *Robot Intelligence Technology and Applications 2: Results from the 2nd International Conference on Robot Intelligence Technology and Applications*. Cham: Springer International Publishing, pp. 593–600.
- Saidi, R., Ncir, W. B. and Essoussi, N. (2018) 'Feature Selection Using Genetic

- Algorithm for Big Data’, *Advances in Intelligent Systems and Computing*, 723, pp. 352–361.
- Seijo-Pardo, B., Bolón-Canedo, V. and Alonso-Betanzos, A. (2016) ‘Using a feature selection ensemble on DNA microarray datasets’, *ESANN 2016 - 24th European Symposium on Artificial Neural Networks*, (April), pp. 277–282.
- Sheng, Z., Hailong, C., Chuan, J. and Shaojun, Z. (2015) ‘An adaptive time window method for human activity recognition’, *Canadian Conference on Electrical and Computer Engineering*. IEEE, 2015-June(June), pp. 1188–1192.
- Shoaib, M., Bosch, S., Durmaz Incel, O., Scholten, H. and Havinga, P. J. M. (2014) ‘Fusion of smartphone motion sensors for physical activity recognition’, *Sensors (Switzerland)*, 14(6), pp. 10146–10176.
- Shoaib, M., Bosch, S., Incel, O. D., Scholten, H. and Havinga, P. J. . (2016) ‘Complex human activity recognition using smartphone and wrist-worn motion sensors’, *Sensors (Switzerland)*, 16(4), pp. 1–24.
- Shoaib, M., Bosch, S., Incel, O. D., Scholten, H. and Havinga, P. J. M. (2016) ‘Complex human activity recognition using smartphone and wrist-worn motion sensors’, *Sensors (Switzerland)*, 16(4).
- Shravankumar, B. and Ravi, V. (2015) ‘Text Classification Using Ensemble Features Selection and Data Mining Techniques’, in Panigrahi, B. K., Suganthan, P. N., and Das, S. (eds) *Swarm, Evolutionary, and Memetic Computing*. Cham: Springer International Publishing, pp. 176–186.
- Shreem, S. S., Abdullah, S. and Nazri, M. Z. A. (2014) ‘Hybridising harmony search with a Markov blanket for gene selection problems’, *Information Sciences*. Elsevier Inc., 258, pp. 108–121.
- Singh, S. and Selvakumar, S. (2015) ‘A hybrid feature subset selection by combining filters and genetic algorithm’, in *International Conference on Computing, Communication and Automation, ICCCA 2015*, pp. 283–289.
- Sivapriya, T. R., Kamal, A. R. N. B. and Thangaiah, P. R. J. (2015) ‘Ensemble Merit Merge Feature Selection for Enhanced Multinomial Classification in Alzheimer’s Dementia’, *Computational and Mathematical Methods in Medicine*, 2015.
- Su, X., Tong, H. and Ji, P. (2014) ‘Activity recognition with smartphone sensors’, *Tsinghua Science and Technology*, 19(3), pp. 235–249.
- Su, Y., Shan, S., Chen, X. and Gao, W. (2009) ‘Hierarchical Ensemble of Global and

- Local Classifiers for Face Recognition’, *IEEE Transactions on Image Processing*, 18(8), pp. 1885–1896.
- Thu, N. T. H. and Han, D. S. (2020) ‘Utilization of Postural Transitions in Sensor-based Human Activity Recognition’, *2020 International Conference on Artificial Intelligence in Information and Communication, ICAIIC 2020*. IEEE, pp. 177–181.
- Too, J., Abdullah, A., Mohd Saad, N., Mohd Ali, N. and Tee, W. (2018) ‘A New Competitive Binary Grey Wolf Optimizer to Solve the Feature Selection Problem in EMG Signals Classification’, *Computers*, 7(4), p. 58.
- Too, J., Abdullah, A. R., Saad, N. M. and Tee, W. (2019) ‘EMG feature selection and classification using a Pbest-guide binary particle swarm optimization’, *Computation*, 7(1).
- Tran, B., Zhang, M. and Xue, B. (2016) ‘A PSO based hybrid feature selection algorithm for high-dimensional classification’, *2016 IEEE Congress on Evolutionary Computation (CEC)*, pp. 3801–3808.
- Trost, S. G., Zheng, Y. and Wong, W. K. (2014) ‘Machine learning for activity recognition: Hip versus wrist data’, *Physiological Measurement*, 35(11), pp. 2183–2189.
- Urbanowicz, R. J., Meeker, M., La Cava, W., Olson, R. S. and Moore, J. H. (2018) ‘Relief-based feature selection: Introduction and review’, *Journal of Biomedical Informatics*.
- Uzer, M. S., Yilmaz, N. and Inan, O. (2013) ‘Feature selection method based on artificial bee colony algorithm and support vector machines for medical datasets classification’, *The Scientific World Journal*, 2013.
- Visalakshi, S. and Radha, V. (2017) ‘A hybrid filter and wrapper feature selection approach for detecting contamination in drinking water management system’, *Journal of Engineering Science and Technology*, 12(7), pp. 1819–1832.
- Walse, K. ., Dharaskar, R. . and Thakare, V. . (2016) ‘Performance Evaluation of Classifiers on Activity Recognition for Disasters’, *Proceedings of the Second International Conference on Information and Communication Technology for Competitive Strategies*, 26, pp. 11–19.
- Walse, K. H., Dharaskar, R. V. and Thakare, V. M. (2016) ‘A study of human activity recognition using adaboost classifiers on WISDM dataset’, *IIOAB Journal*, 7(2), pp. 68–76.

- Wang, W., Liu, A. X., Shahzad, M., Ling, K. and Lu, S. (2017) ‘Device-Free Human Activity Recognition Using Commercial WiFi Devices’, *IEEE Journal on Selected Areas in Communications*. IEEE, 35(5), pp. 1118–1131.
- Wang, Y., Cang, S. and Yu, H. (2019) ‘A survey on wearable sensor modality centred human activity recognition in health care’, *Expert Systems with Applications*. Elsevier Ltd, 137, pp. 167–190.
- Wang, Y., Vasilakos, A., Jin, Q. and Zhu, H. (2017) ‘Smartphone-Based Human Activity Recognition’, *Device-to-Device based Proximity Service*, (2012), pp. 199–239.
- Wang, Z. and Meng, F. (2018) ‘An overview of human activity recognition based on smartphone’.
- Wang, Z., Meng, F., Yuan, G., Yan, Q. and Xia, S. (2018) ‘An overview of human activity recognition based on smartphone’, *Sensor Review*.
- Wosiak, A. and Dziomdziora, A. (2015) ‘Feature selection and classification pairwise combinations for high-dimensional tumour biomedical datasets’, *Schedae Informaticae*, 24, pp. 53–62.
- Yamansavaşçılar, B. and Güvensan, M. A. (2016) ‘Activity Recognition on Smartphones: Efficient Sampling Rates and Window Sizes’, *2016 IEEE International Conference on Pervasive Computing and Communication Workshops, PerCom Workshops 2016*, (PerMoby).
- Yusup, N., Zain, A. M. and Latib, A. A. (2019) ‘A review of Harmony Search algorithm-based feature selection method for classification’, *Journal of Physics: Conference Series*, 1192(1).
- Zainuddin, Z., Lai, K. H. and Ong, P. (2016) ‘An enhanced harmony search based algorithm for feature selection: Applications in epileptic seizure detection and prediction’, *Computers and Electrical Engineering*. Elsevier Ltd, 53, pp. 143–162.
- Zainudin, M. N. S., Sulaiman, M. N., Mustapha, N. and Perumal, T. (2017a) ‘Recognizing stationary and locomotion activities using combinational of spectral analysis with statistical descriptors features’, *AIP Conference Proceedings*, 1891(October).
- Zainudin, M. N. S., Sulaiman, M. N., Mustapha, N. and Perumal, T. (2017b) ‘Recognizing stationary and locomotion activities using combinational of spectral analysis with statistical descriptors features’, *AIP Conference*

Proceedings, 1891(October).

- Zdravevski, E., Lameski, P., Trajkovik, V., Kulakov, A., Chorbev, I., Goleva, R., Pombo, N. and Garcia, N. (2017) ‘Improving Activity Recognition Accuracy in Ambient-Assisted Living Systems by Automated Feature Engineering’, *IEEE Access*, 5, pp. 5262–5280.
- Zheng, Y. (2015) ‘Human Activity Recognition Based on the Hierarchical Feature’, *IEEE International Conference on Information and Automation*, 2015(July), pp. 219–224.
- Zubair, M., Song, K. and Yoon, C. (2016) ‘Human activity recognition using wearable sensors’, *Advances in Intelligent Systems and Computing*. IEEE, 1177, pp. 527–538.

LIST OF PUBLICATIONS

Conference Proceedings

- Yusup, N., Zain, AM., Latib, AA (2018), A review of Harmony Search algorithm-based feature selection method for classification, Tthe 2nd International Conference On Data And Information Science (ICoDIS), IOP Journal of Physics: Conference Series (Indexed by SCOPUS) – (Accepted)
- Yusup, N. and Mohd Zain, A. (2019) ‘The Performance Review of mRMR for Gene Selection and Classification of DNA Microarrays’, *IOP Conference Series: Materials Science and Engineering*, 551, p. 012040. (Indexed by SCOPUS) – (Accepted)
- Fatin Liyana Mohd Rosely, N., Mohd Zain, A., Yusoff, Y. and Yusup, N. (2019) ‘A Brief Conceptual View on Classification Using Support Vector Machine’, *IOP Conference Series: Materials Science and Engineering*, 551, p. 012035. (Indexed by SCOPUS) – (Accepted)
- Yusup, N., Zain, A.M, Mohd Rosely NL, Yusuf, SM (2018), Feature Selection with Harmony Search for Classification: A Review, The 3rd International Conference of Computer, Environment, Social Science, Health Science, Agriculture and Technology (ICEST) – (Accepted and presented)