[Sleep Apnea Detection Using Multi-Bio Signals and Machine Learning]

by [Xilin Li]

Thesis submitted in fulfilment of the requirements for the degree of

[Doctor of Philosophy]

under the supervision of [Dr Ling, Steve Sai Ho]

University of Technology Sydney
Faculty of [Engineering and Information Technology]

[August 2020]
Declaration

I, Xilin Li declare that this thesis, is submitted in fulfilment of the requirements for the award of Doctor of Philosophy, in the Faculty of Engineering and Information Technology at the University of Technology Sydney.

This thesis is wholly my own work unless otherwise referenced or acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

This document has not been submitted for qualifications at any other academic institution.

*If applicable, the above statement must be replaced with the collaborative doctoral degree statement (see below).

*If applicable, the Indigenous Cultural and Intellectual Property (ICIP) statement must be added (see below).

This research is supported by the Australian Government Research Training Program.

Production Note: Signature: Signature removed prior to publication.

Date: 02/08/2021
Acknowledgements

I want to express my sincere gratitude to my Principal Supervisor Dr Ling, Steve Sai Ho, for his continual support, guidance, help, and encouragement during my PhD period. Dr Ling gives me solid support in the topic of Sleep Apnea Detection Using Multi-Bio Signals and Machine Learning, and always provides brilliant insights into my research works. It is my honour to have such a supervisor who always inspires me to achieve higher targets. His conscientious and meticulous attitude on research has a significant influence on my work.

I want to express my sincere gratitude to my Co-supervisor Associate Professor Su, (Steven) Steven Weidong, for their precious comments and suggestions on my research. I also would like to demonstrate my sincere gratitude to my External Supervisor Al-Ani, (Ahmed) Ahmed. He has brought me into the topic of Sleep Apnea Detection and given me his kind support in the first year of my PhD degree.

Special thanks my friends in the CHT research group, in particular, Lin Ye, Wenhui Chen, Haimin Zhang, Zhiyuan Shi, Zhichao Sheng, Kairui Guo, Hanjie Wu, Lingmeng Li for their selfless help. Working together with them will be a good memory.

My deepest gratitude goes to my parents and my girlfriend for their immeasurable support and encouragement throughout my graduate studies.
Abstract

One of the most common types of breathing diseases is obstructive sleep apnea (OSA). OSA affects about 2% - 5% of the total human population. Polysomnography (PSG) is considered the gold standard and records multichannel bio-signals throughout one entire night, including electrocardiogram (ECG), electroencephalogram (EEG), electrooculogram (EOG), electromyogram (EMG), airflow (AF), abdominal and thoracic efforts (AB and TH), body position, snore, oxygen saturation (SaO₂). Patients may not be able to sleep well because an OSA test is always performed in a sleep laboratory or a hospital, and electrodes are placed on the patient’s skin. On the other hand, manual detection of OSA is time-consuming and costly because sleep experts need to monitor and review the overnight PSG signals. To solve these problems, an automatic OSA detection system with high accuracy is desirable.

In this study, the main objective is to construct an OSA monitoring system with high accuracy that is based on multi-domain features from multiple bio-signals using machine learning methods. To achieve this main objective, there are four main contributions presented as follows: 1) feature extraction, 2) feature selection, 3) classification for OSA, and 4) the performance of this monitoring system with the time-window method. Firstly, time-domain, frequency-domain, and non-linear analysis algorithms are used to extract features from multiple bio-signals, such as SaO₂, AF, AB, TH, and ECG recordings. The bio-signals have different bio-patterns between apnea and
normal events, and the features are also related to the apneic bio-physiological patterns. Secondly, a hybrid two-stage feature selection is proposed to select the significant features. Statistical analyses (the first stage) are used to determine independent and significant features, and selected features are put into the feature subset. In the second stage, machine learning methods are utilised to confirm the final feature subset. Thirdly, a multi-errors-reduction (MER) classification system is constructed to classify apnea or normal episodes. Considering the different machine learning algorithms, the structure of the MER system is the stacking method, and it consists of weak learners (boosting methods) and a meta-learner (an artificial neural network (ANN)). Fourthly, the proposed monitoring system is able to provide good performance in the 60-second time-window segmentation method, which means that this monitoring system has the potential for implementation using real data.

In this thesis, the frequently used database was the Sleep Heart Health Study (SHHS) database. This dataset provides over 1,500 patients’ PSG signals and each respiratory event and its duration is labelled by a human expert, which is able to evaluate the stability of the proposed monitoring system. According to the duration of each event, multi-domain feature extraction methods are used to provide 66 kinds of features from ECG, SaO₂, AF, TH, and AB signals. The two-stage feature selection procedure then selects 19 kinds of significant features and determines the final feature subset. These 19 features are extracted from ECG, SaO₂, and AB signals. Compared with different machine learning methods, the support vector machine (SVM) method has better performance than other classifiers (66.54% specificity, 97.05% sensitivity, and 81.68% accuracy). It has been found that the SVM method show poor performance when it predicts normal events in less than 60 seconds. Thus, in order to enhance classification results, the 60-second time-window segmentation method is used to improve classification performance and evaluate the potential for implementation.
with real data. Using this time-window method, 48 selected kinds of features, which were extracted from five kinds of bio-signals, obtained better performance (sensitivity = 91.94%, specificity = 89.00%, accuracy = 90.71%). The MER classification system is developed to enhance performance. The structure of the MER system is the stacking method. To construct the MER system, four boosting methods (Gradient Boosting, CatBoost, Light GBM, and XGBoost) are base learners, and the meta-learner is the ANN with one input layer, one hidden layer with 4 knots, and one output layer. This MER classification system is able to hold better performance (sensitivity = 96.37%, specificity = 90.83%, accuracy = 94.66%) compared to existing studies.
Publications

The contents of this thesis are based on the following papers that have been published, accepted, or submitted to peer-reviewed journals and conferences.

Journal Papers:


Conference Papers:

Table of Contents

List of Figures

List of Tables

Nomenclature

1 Introduction ............................................. 1

1.1 Background ............................................ 2
1.2 Motivation of Thesis ................................. 15
1.3 Objectives and Contributions ....................... 16
   1.3.1 Multi-Domain Feature Extraction Algorithms .... 17
   1.3.2 Two-Stage Feature Selection Method ............. 19
   1.3.3 Automatic OSA Monitoring System Based on Time-Window Method ..................................................... 21
   1.3.4 Multi-Errors-Reduction Classification System . . 21
1.4 Structure of Thesis .................................... 22

2 Literature Review ........................................ 25
# Table of Contents

2.1 An Introduction to Automatic OSA Detection Using Bio-Signals . . . . 25
   2.1.1 Automatic Detection with Oxygen Saturation Signals . . . . . . 26
   2.1.2 Automatic Detection with Airflow Signals . . . . . . . . . . . . 27
   2.1.3 Automatic Detection with Respiratory Signals . . . . . . . . . . 29
   2.1.4 Automatic Detection with ECG Signals . . . . . . . . . . . . . . 30
   2.1.5 Automatic Detection with EEG Signals . . . . . . . . . . . . . . 32
   2.1.6 Automatic Detection with Multiple Bio-Signals . . . . . . . . . 33

2.2 An Introduction to Statistical Analyses . . . . . . . . . . . . . . . . 39
   2.2.1 An Introduction to Analysis of Variance . . . . . . . . . . . . . 39
   2.2.2 An Introduction to Rank-Sum Test . . . . . . . . . . . . . . . . 40

2.3 A Brief Introduction to Machine Learning Algorithms . . . . . . . . . . 41
   2.3.1 Support Vector Machine . . . . . . . . . . . . . . . . . . . . . . 41
   2.3.2 Decision Tree Algorithm . . . . . . . . . . . . . . . . . . . . . . 42
   2.3.3 K-Nearest Neighbors Algorithm . . . . . . . . . . . . . . . . . 44
   2.3.4 Artificial Neural Networks . . . . . . . . . . . . . . . . . . . . . 46
   2.3.5 Ensemble Algorithm . . . . . . . . . . . . . . . . . . . . . . . . 48
      2.3.5.1 Random Forest Algorithm . . . . . . . . . . . . . . . . 49
      2.3.5.2 Boosting Algorithm . . . . . . . . . . . . . . . . . . . 50
      2.3.5.3 Stacking Algorithm . . . . . . . . . . . . . . . . . . . . 51

2.4 Summary . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 52

3 Multi-Domain Feature Extraction Methods . . . . . . . . . . . . . . . . . 55
Table of Contents

3.1 Introduction ................................................. 55

3.2 Apneic Bio-Physiological Patterns and Feature Extraction Methods . . 57
   3.2.1 Apneic Patterns and Feature Extraction from SaO$_2$ Signals . . . 57
   3.2.2 Apneic Patterns and Feature Extraction from Airflow Signals . . 60
   3.2.3 Apneic Patterns and Feature Extraction from Respiratory Signals 63
   3.2.4 Apneic Patterns and Feature Extraction from ECG Signals . . . 66

3.3 Feature Extraction Methods on Another Bio-Signals .................. 77
   3.3.1 Multi-Domain Features from EEG Signals .................... 78
   3.3.2 Multi-Domain Features from Chin EMG Signals ............... 79
   3.3.3 Experiments and Results .................................. 81

3.4 Summary ......................................................... 81

4 Feature Selection Method ........................................ 83

4.1 Introduction .................................................... 83

4.2 Two-Stage Feature Selection Method ............................ 84
   4.2.1 Statistics Analysis ........................................ 85
   4.2.2 Support Vector Machine Selection .......................... 87

4.3 Experiments, Results, and Discussion .......................... 91
   4.3.1 Collection of Sleep Studies ................................. 91
   4.3.2 Feature Extraction Experiments and Results ............... 92
   4.3.3 Feature Selection Experiments and Results ............... 92
   4.3.4 Discussion ................................................. 96
4.4 Summary ............................................................. 99

5 Classification Based on Duration Parameter and Evaluation of The
Time-Window Method .............................................. 101

5.1 Introduction ......................................................... 101

5.2 Classification Based on Duration Parameter ..................... 102

5.2.1 Experiments and Results .................................... 102

5.2.2 Discussion ....................................................... 104

5.3 Evaluation of The 60-Second Time-Window Algorithm .......... 106

5.3.1 Feature Extraction ............................................ 106

5.3.2 Feature Selection .............................................. 107

5.3.3 Classification .................................................. 110

5.3.4 Discussion ....................................................... 111

5.4 Summary ............................................................ 113

6 Obstructive Sleep Apnea Classification Using The Multi-Errors-Reduction
(MER) System Based on The 60-Second Time-Window Method ........ 115

6.1 Introduction ......................................................... 115

6.2 Classification by Machine Learning Methods .................... 117

6.2.1 Feature Extraction and Feature Selection .................... 117

6.2.2 Classification Based on Ranking Order ...................... 121

6.3 Multi-Errors-Reduction Classification System .................. 123

6.3.1 Methodology .................................................... 124
# Table of Contents

6.3.2 Experiments, Results and Discussion .......................... 127  
6.3.3 Discussion ....................................................... 131  
6.4 Summary ............................................................ 132  

7 Conclusion and Future Work .......................................... 135  
7.1 Conclusion .......................................................... 135  
7.2 Future Work ....................................................... 138  

References ................................................................. 141
List of Figures

1.1 The patient’s airway gets blocked because of obstructive sleep apnea (Source: mayoclinic.org) ........................................ 3

1.2 Example of Polysomnography, including EEG, EOG, EMG, abdominal, thoracic, airflow, SaO$_2$, heart rate, ECG signals (Source: physionet.org) 5

1.3 Example of a pulse oximetry attaches to a finger (Source: theprint.in) . 6

1.4 Example of 300-second session shows sudden drops, fast recoveries, and typical saw-tooth morphology of the oximetry curve . ................. 6

1.5 Example of the airflow sensor placed under a patient’s nose (Source: jcsm.aasm.org) ....................................................... 8

1.6 Example of airflow, abdominal, thoracic, SaO$_2$ signals, and two respiratory efforts are related with the desaturation recordings (airflow and SaO$_2$ signals) [1] ....................................................... 8

1.7 Example of thoracic and abdominal bands (Source: link.springer.com) . 9

1.8 Example of ECG signal shows P, Q, R, S, T waves, QRS complex and RR interval (Source: medictests.com) ............................ 10
1.9 Difference between apnea and normal windows in ECG signal (the top wave is the normal ECG window and the bottom wave is the abnormal ECG window) .......................................................... 10

1.10 10-20 system EEG placement and electrode distances (Source: ieeexplore.ieee.org) .......................................................................................................................... 12

1.11 the EEG signal and $\delta$, $\theta$, $\alpha$, and $\beta$ component waveforms (Source: frontiersin.org) ......................................................................................................................... 13

1.12 EOG electrode placement and the two EOG electrodes are placed below the outer canthus of each eye, and EMG electrode placement is used to measure the changes of the geniohyoid muscle (Source: cvgclinical.co.za) 14

1.13 Basic components of the OSA automatic detection progress ............ 16

2.1 An example of decision tree (Source: matrix.squiz.net) .................... 43

2.2 The basic components of an artificial neural network ......................... 45

2.3 The basic components of the boosting method ................................. 51

2.4 The basic components of the stacking method ................................. 52

3.1 The oxygen saturation difference in time domain between one apnea event and one normal event (the top wave is the oxygen saturation in the normal event and the bottom wave is the oxygen saturation in the abnormal event) ............................................................................. 58

3.2 The oxygen saturation difference in frequency domain between one apnea event and one normal event (the top wave is the frequency wave of the normal event and the bottom wave is the wave of the abnormal event) 58
List of Figures

3.3 The airflow difference in time domain between one apnea event and one normal event (the top wave is the airflow signal in the normal event and the bottom wave is the airflow signal in the abnormal event) .......................... 60

3.4 The airflow difference in frequency domain between one apnea event and one normal event (the top wave is the frequency wave of the normal event and the bottom wave is the wave of the abnormal event) .......................... 61

3.5 The abdominal difference in time domain between one apnea event and one normal event (the top wave is the abdominal signal in the normal event and the bottom wave is the abdominal signal in the abnormal event) ............................................................................................................. 63

3.6 The thoracic difference in time domain between one apnea event and one normal event (the top wave is the thoracic signal in the normal event and the bottom wave is the thoracic signal in the abnormal event) .......................... 64

3.7 The one-minute ECG signal (the top) and its QRS complex (the bottom) .......................... 66

3.8 The one-minute ECG signal (the top) and its RR interval series (the bottom) ..................................................................................................................................................... 66

3.9 The R amplitude difference between one apnea event and one normal event (the top wave is the normal R amplitude series and the bottom wave is the abnormal R amplitude series) .................................................................................................................. 68

3.10 The RR interval difference between one apnea event and one normal event (the top wave is the normal RR interval series and the bottom wave is the abnormal RR interval series) .................................................................................................................................................. 68

3.11 The ECG derivation respiratory (EDR) difference between one apnea event and one normal event (the top wave is the normal EDR signal and the bottom wave is the abnormal EDR signal) .................................................................................................................................................. 69
3.12 The ECG difference in frequency domain between one apnea event and one normal event (the top wave is the ECG frequency wave of the normal event and the bottom wave is the wave of the abnormal event). 71

3.13 The EDR difference in frequency domain between one apnea event and one normal event (the top wave is the EDR frequency wave of the normal event and the bottom wave is the wave of the abnormal event). 71

3.14 The RR interval difference in frequency domain between one apnea event and one normal event (the top wave is the RR interval frequency wave of the normal event and the bottom wave is the wave of the abnormal event) 72

3.15 Example of ECG, EMG, thoracic, and abdominal signals from the DREAMS dataset 80

3.16 Example of EEG, SaO\textsubscript{2}, and airflow signals from the DREAMS dataset 81

4.1 Two-stage feature selection procedure, including Stage 1: the statistical analysis stage and Stage 2: the final feature subset selection stage using machine learning methods 86

4.2 The flowchart of calculating the importance level of each feature 88

4.3 The hill-climbing method based on the descending order list 90

4.4 One apnea event duration in SaO\textsubscript{2} signal, and the duration is between the apnea start line and the apnea end line 91

4.5 Example of PSG including ECG, abdominal and thoracic signals, airflow and oxygen saturation signals 92
6.1 The hill-climbing method based on the descending order list in the 60-second time-window experiment ................................. 120
6.2 The iteration performance of the hill-climbing algorithm based on the descending order list using the SVM model in the 60-second time-window segmentation method ........................................ 122
6.3 Basic components of the multi-errors-reduction classification system .............................. 124
6.4 All steps of each weak learner in stacking combination method ........................................ 125
6.5 Training process of an ANN in the multi-errors-reduction classification system .................. 128
6.6 Prediction process of the trained ANN in the multi-errors-reduction classification system ........ 129
List of Tables

2.1 Studies to automatically detect OSA based on bio-signals . . . . . . . . 36
2.2 Summary of apneic event detection approaches using single one bio-signal 38
2.3 Summary of apneic event detection approaches using multi-bio signals . 39
2.4 List of kernel functions used in support vector machine models . . . . . 42

3.1 Multi-domain features from SaO$_2$ signals (The abbreviations are shown
and detail methods are shown in Section 3.2.1) . . . . . . . . . . . . . . 59
3.2 Multi-domain features from airflow signals (The abbreviations are shown
and detail methods are shown in Section 3.2.2) . . . . . . . . . . . . . . 62
3.3 Multi-domain features from abdominal signals (The abbreviations are
shown and detail methods are shown in Section 3.2.3) . . . . . . . . . . 65
3.4 Multi-domain features from thoracic signals (The abbreviations are
shown and detail methods are shown in Section 3.2.3) . . . . . . . . . . 65
3.5 Multi-domain features from ECG signals (The abbreviations are shown
and detail methods are shown in Section 3.2.4) . . . . . . . . . . . . . . 70
3.6 Features extracted from SaO₂, airflow, abdominal signal, thoracic signal, and ECG signals (No.1 - No.12 from SaO₂ signals, No.13 - No.22 from airflow signals, No.23 - No.28 from abdominal signals, No.29 - No.34 from thoracic signals, and feature No.35 - No.66 from ECG signals, four dashed lines divided the table into five parts according to different kinds of signals) .................................................. 73

3.7 Features extracted from EEG signals (The abbreviations are shown and detail methods are shown in Section 3.3.1) .................................................. 79

3.8 Features extracted from EMG signals (The abbreviations are shown and detail methods are shown in Section 3.3.2) .................................................. 80

4.1 Feature selection results in the experiment with duration using statistics analysis in the SHHS dataset (No.1 - No.12 from oxygen saturation signals, No.13 - No.22 from airflow signals, No.23 - No.28 from abdominal signals, No.29 - No.34 from thoracic signals, feature No.35 - No.66 from ECG signals, four dashed lines divided the table into five parts according to different kinds of signals) .................................................. 94

4.2 Feature classes via the distribution of λ_{feature} .................................................. 95

4.3 Sensitivity (%), specificity (%) and accuracy (%) based on the hill-climbing method using SVM models with different kernel functions and parameters in the experiment with duration .................................................. 97

4.4 AUC (%) computed on SVM models with kernels and parameters using Class ABC and ABCD in the experiment with duration .................................................. 98

5.1 Performance of each classifier with the 19 selected features from ECG, SaO₂ and abdominal signals .................................................. 104
List of Tables

5.2 The duration and its count of each normal event in Patient No. 1537 105
5.3 The duration and its count of each normal event in Patient No. 1490 105
5.4 Features extracted from SaO₂, airflow, abdominal, thoracic, and ECG signals based on the 60-second time-window segmentation method (No.1 - No.12 from SaO₂ signals, No.13 - No.22 from airflow signals, No.23 - No.28 from abdominal signals, No.29 - No.34 from thoracic signals, and feature No.35 - No.66 from ECG signals, four dashed lines divided the table into five parts according to different kinds of signals) 108
5.5 Feature selection results in the experiment with the 60-second time-window method using statistics analysis (No.1 - No.12 from SaO₂ signals, No.13 - No.22 from airflow signals, No.23 - No.28 from abdominal signals, No.29 - No.34 from thoracic signals, and feature No.35 - No.66 from ECG signals, four dashed lines divided the table into five parts according to different kinds of signals) 109
5.6 Feature classes via the distribution of λ_{feature} in the experiment with the 60-second time-window method 110
5.7 Sensitivity (%), specificity (%), accuracy (%), and AUC (%) based on the hill-climbing method using the SVM model (RBF, σ = 25, R = 0.2) with six selected features from SaO₂ signals 110
5.8 Performance of each classifier with the 66 selected features in the experiment with the 60-second time-window method 111
5.9 Comparison the SVM performance of the duration method and the SVM performance of the 60-second time-window method 112
<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.1</td>
<td>The descending order of 66 features according to $\lambda_{\text{feature}}$ in the 60-second time-window experiment (the bio-signal provides the feature)</td>
</tr>
<tr>
<td>6.2</td>
<td>Performance of each classifier with the 48 selected features in the 60-second experiment</td>
</tr>
<tr>
<td>6.3</td>
<td>Performance of four weak classifiers and the multi-errors-reduction classifier with the 48 selected features based on the 60-second time-window segmentation method</td>
</tr>
<tr>
<td>6.4</td>
<td>Performance of different classifiers and the multi-errors-reduction classifier with the 48 selected features based on the 60-second time-window segmentation method</td>
</tr>
<tr>
<td>6.5</td>
<td>Performance of each classifier and the multi-errors-reduction classifier with the 19 selected features from ECG, $\text{SaO}_2$, and abdominal signals, based on the duration parameter</td>
</tr>
</tbody>
</table>
Nomenclature

Symbols

$\lambda_{\text{feature}}$ the importance of a feature

d the result of statistical analyses

$p$-value the result of statistical analyses

$\alpha$ alpha wave

$\beta$ beta wave

$\delta$ delta wave

$N$ the importance of a feature

$\theta$ theta wave

Acronyms / Abbreviations

$\sigma$ Width of the Radial basis function kernel

d Degree of the Polynomial kernel

$FN$ False Negative

$FP$ False Positive
<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N$</td>
<td>Condition Negative</td>
</tr>
<tr>
<td>$P$</td>
<td>Condition Positive</td>
</tr>
<tr>
<td>$R$</td>
<td>Regularization parameter of the support vector machine method</td>
</tr>
<tr>
<td>$TN$</td>
<td>True Negative</td>
</tr>
<tr>
<td>$TP$</td>
<td>True Positive</td>
</tr>
<tr>
<td>AB</td>
<td>Abdominal</td>
</tr>
<tr>
<td>AF</td>
<td>Airflow</td>
</tr>
<tr>
<td>AHI</td>
<td>Apnea-Hypopnea Index</td>
</tr>
<tr>
<td>ANN</td>
<td>Artificial Neural Network</td>
</tr>
<tr>
<td>ANOVA</td>
<td>Analysis of Variance</td>
</tr>
<tr>
<td>AUC</td>
<td>Area Under The ROC Curve</td>
</tr>
<tr>
<td>ECG</td>
<td>Electrocardiogram</td>
</tr>
<tr>
<td>EDR</td>
<td>ECG-Derived Respiration</td>
</tr>
<tr>
<td>EEG</td>
<td>Electroencephalogram</td>
</tr>
<tr>
<td>EMG</td>
<td>Electromyogram</td>
</tr>
<tr>
<td>EOG</td>
<td>Electrooculogram</td>
</tr>
<tr>
<td>HRV</td>
<td>Heart Rate Variability</td>
</tr>
<tr>
<td>kNN</td>
<td>k-nearest neighbour algorithm</td>
</tr>
<tr>
<td>Light GBM</td>
<td>Light Gradient Boosting</td>
</tr>
</tbody>
</table>
Nomenclature

MER Multi-Errors-Reduction

OSA Obstructive Sleep Apnea

PCA Principal Component Analysis

PSD Power Spectral Density

PSG Polysomnography

RBF Radial Basis Function kernel

SaO$_2$ Oxygen Saturation

Sen Sensitivity

SHHS Sleep Heart Health Study

Spe specificity

SVM Support Vector Machine

TH Thoracic

WSD Wavelet Spectral Density