

**ADVANCED DEEP LEARNING APPROACHES FOR BIOSIGNALS
APPLICATIONS**

By Marwa Farouk Ibrahim Ibrahim

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

PhD in the School of Biomedical Engineering, Faculty of
Engineering and Information Technology
The University of Technology Sydney (UTS)

January 2019

Certificate of Original Authorship

I, Marwa Farouk Ibrahim, certify that the work in this thesis has not previously been submitted for a degree nor has it been submitted as part of requirements for a degree except as fully acknowledged within the text.

I also certify that the thesis has been written by me. Any help that I have received in my research work and the preparation of the thesis itself has been acknowledged.

In addition, I certify that all information sources and literature used are indicated in the thesis.

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

Signature of Student: Production Note:
Signature removed prior to publication.

Date: 11 January 2019

Acknowledgment

I would like to express my deep gratitude and extend my warmest and most sincere thanks to my co-supervisor **Associate Professor Robert Fitch** for all his continuous support, efforts, endless help, guidance, understanding and discussions.

I would like to thank **Dr Nham Tran** for his generous support which otherwise I will not be able to finish my thesis.

I would like also to thank **Associate Professor Gyorgy Hutvagner** for his powerful decision that helped me finish my thesis.

I would like to extend further my genuine appreciations to **Dr Ben Rodanski** for giving me the chance of working as a casual academic at UTS.

I would also like to thank **Dr. Khairul Anam** who generously supported me with the EMG datasets and valuable discussions.

I would like to pass my thanks to **Associate Professor Adel Al- Jumaily** for the datasets and resources that he provided me.

Last but not least, I would like to express my special gratitude to my beloved father **Farouk**, wonderful mother **Faten**, lovely brothers **Hesham** and **Emad** and my mother-in-law **Nour El Hoda** for their continuous support that motivated me to finish my thesis.

A very special thanks goes to my beloved husband, **Tawfek Mahmoud**, for his endless support, love, encouragement, motivation, and patience during my entire years of study.

Table of Contents

Certificate of Original Authorship	i
Acknowledgment	ii
List of Figures	vii
List of Tables	x
List of Abbreviations	xi
Abstract	xvi
CHAPTER 1	1
Problem background	1
1.1 Background-global disability rates	1
1.2 Hand disability	4
1.3 Summary of challenges	7
1.4 Objectives	8
1.5 Contribution to the doctoral thesis	9
1.6 Organization of the thesis	12
1.7 Publications outcome of the doctoral research	12
CHAPTER 2	14
Literature review	14
2.1 Hand Anatomy	14
2.1.1 Bones	14
2.1.2 Skin	16
2.1.3 Nerves	17
2.1.4 Muscles and Tendons	19
2.1.5 Joints	21
2.2 Biosignals	22
2.2A Electromyography	23
2.3 Finger Movement Classification	31
2.4 Biosignal Pattern Recognition	59
2.4.1 Pre-Processing	59
2.4.2 Feature extraction	61
2.4.3 Feature reduction	66
2.4.4 Classification	70
2.4.3 Post-processing	82
2.5 Summary	83
CHAPTER 3	85
Review on feature learning, deep learning and data representation	85

3.1 Feature selection	85
3.1.1 Subset Generation	86
3.1.2 Subset Evaluation.....	87
3.1.3 Stopping criterion	89
3.1.4 Result Validation.....	90
3.2 Implementation of feature selection for biosignals.....	90
3.3 Feature Learning	94
3.3.1 Supervised learning	95
3.3.2 Unsupervised learning	96
3.4 Deep Learning.....	99
3.5 Artificial neural networks	105
3.5.1 Neural networks architecture.....	107
3.5.2 Neural networks learning techniques	109
3.6 Signal representation.....	115
• Spectrogram.....	117
• Wavelet.....	119
• Wavelet packet	120
3.7 Data Augmentation.....	122
3.8 Tensor	124
3.9 Classifier fusion	126
• Dynamic classifier selection	127
• Classifier structuring and grouping.....	127
• Hierarchical mixture of experts	128
3.10 Summary	128
CHAPTER 4	130
Evaluation of Auto Encoder Deep learning for biosignal.....	130
4.1 Implementation of Auto Encoder Deep learning for biosignal.....	131
4.1.1 Previous work	132
4.2 Experiments and Results.....	134
4.2.1 Data Acquisition	134
4.2.2 Signal Representation	136
4.2.3 Results.....	136
4.2.4 Comparison between the proposed model with other researchers' achievements	141
4.3 Discussion.....	144
4.4 Summary.....	146
CHAPTER 5	148
Evaluation of PCA and ICA Feature Learning for Biosignals.....	148

5.1 Implementation of PCA and ICA feature learning for biosignal	149
5.1.1 <i>Independent component analysis</i>	150
5.1.2 <i>Manhattan index</i>	150
5.2 Experiments and Results	151
5.2.1 <i>Results</i>	151
5.2.2 <i>Comparison between the proposed model with other researchers' achievements</i>	155
5.3 Discussion	157
5.4 Summary	158
CHAPTER 6	160
Evaluation of Self Organising Map Deep Learning for Biosignals	160
6.1 Implementation of Self Organising Map Deep learning for biosignal	161
6.1.1 <i>Previous work</i>	167
6.2 Experiments and Results	169
6.2.1 <i>Results</i>	169
6.2.2 <i>Comparison between the proposed model with other researchers' achievements</i>	173
6.3 Discussion	175
6.4 Summary	176
CHAPTER 7	178
Evaluation of Augmented Data with Tensor Representation for Deep Learning for Biosignals	178
7.1 Implementation of augmented data with tensor representation for deep learning for biosignals	179
7.1.1 <i>Previous work</i>	180
7.2 Experiments and Results	182
7.2.1 <i>Data Acquisition</i>	182
7.2.2 <i>Results</i>	184
7.2.3 <i>Comparison between the proposed model with other researchers' achievements</i>	187
7.3 Discussion	190
7.4 Summary	190
CHAPTER 8	192
Summary and Conclusion	192
8.1 Thesis Summary	192
8.2 Recommendations for future research	197
8.2.1 <i>Improving deep learning techniques</i>	197
8.2.2 <i>Data fidelity and precision</i>	198
8.2.3 <i>Developing the evaluation techniques</i>	199
8.2.4 <i>Enhancing classifier performance</i>	199
8.2.5 <i>Evolving the data acquisition approaches</i>	199
8.2.6 <i>Targeting more classes</i>	200

8.2.7 <i>Reducing the simulation time</i>	200
8.2.8 <i>Applying outstanding models in real applications</i>	200
8.3 Conclusion	201
<i>References</i>	203

List of Figures

Figure 1. 1: Disability percentage between boys and girls in four countries [10]	3
Figure 1. 2: Disability percentage for Low-income and high-income countries [4].....	4
Figure 1. 3: Disability percentage by age (Canadian Study) [13].....	4
Figure 1. 4: The suggested innovative prosthetic hand (i-limb) [14].....	5
Figure 2. 1: Bones groups in hand [31].....	15
Figure 2. 2: Hand Bones dorsal view [31]	15
Figure 2. 3: Hand Bones Palmar view [31].....	15
Figure 2. 4: Hand skin for palm and dorsum [31].....	17
Figure 2. 5: EMG signal pattern [48].....	23
Figure 2. 6: EMG signal pattern in black and its energy in red [49].....	24
Figure 2. 7: Motor unit [48].....	24
Figure 2. 8: Depolarisation and polarisation of muscle membrane excitation [48]	25
Figure 2. 9: The Action potential [48]	26
Figure 2. 10: The superimposition of motor unit action potential [48].....	26
Figure 2. 11: The superimposition of multiple motor unit with different firing frequencies [48].....	27
Figure 2. 12: The raw EMG for three contractions recorded from the biceps brachii muscle [48]	27
Figure 2. 13: The production and recording of EMG [18].....	29
Figure 2. 14: EMG Recording stages and tools [53].....	29
Figure 2. 15: EGG interference with EMG [48]	31
Figure 2. 16: EMG different applications areas.....	31
Figure 2. 17: Error percentage for classifying five-finger movements for six subjects [61]	34
Figure 2. 18: EMG signal recording via four electrodes [61]	34
Figure 2. 19: Normalised training and testing distribution [61].....	35
Figure 2. 20: ellipsoidal delimitation distribution for trimmed training and testing set [61].....	36
Figure 2. 21: A. Front view for electrodes allocation on intact subjects [22].....	37
Figure 2. 22: Suggested classification system with four combined schemes [22].....	38
Figure 2. 23: Implementation of error for number of classes with four schemes [22].....	38
Figure 2. 24: Classification error for proposed model versus number of channels [22].....	38
Figure 2. 25: The suggested scheme block diagram [66].....	39
Figure 2. 26: Accuracy versus number of neurons for all participants [66]	41
Figure 2. 27: Accuracy versus number of neurons subject based [66]	41
Figure 2. 28: Block Diagram for the suggested pattern recognition model [77]	42
Figure 2. 29: Block Diagram for classifying ten finger movements by using AW-ELM [79]	43
Figure 2. 30: Suggested model performance for different feature reduction algorithms versus the number of classes [80].....	44
Figure 2. 31: Suggested model accuracy value versus number of nodes [80]	45
Figure 2. 32: Suggested model accuracy value versus regression coefficient α [80]	45
Figure 2. 33: Average accuracy for all subjects and individually for four classifiers [81].....	47
Figure 2. 34: Average accuracy against the number of hidden nodes during offline classification [81].....	47
Figure 2. 35: Average accuracy for subject 1 for three consecutive days [81]	48
Figure 2. 36: Electrodes allocation for healthy and amputee participants [22].....	48
Figure 2. 37: Block diagram for the suggested system [82]	49
Figure 2. 38: Average accuracy versus number of channels [82]	50
Figure 2. 39: Average accuracy against window length [82].....	51
Figure 2. 40: Processing time versus window length [82]	51
Figure 2. 41: Electrodes allocation on right forearm in two positions (a) and (b) [23]	52
Figure 2. 42: Classification accuracy for each dimension reduction for different classes [23]	52

Figure 2. 43: Electrodes placement in two positions (a) and (b) [83].....	53
Figure 2. 44: Targeted ten finger movements [83]	53
Figure 2. 45: Classification error attained by implementing different dimension reduction along with three classifiers [83].....	54
Figure 2. 46: Classification error for three different classifiers with different window sizes [20].....	55
Figure 2. 47: Electrodes allocation at interior and posterior positions in addition to the utilised hardware in collecting the data [86].....	56
Figure 2. 48: Targeted fifteen finger gestures [86]	56
Figure 2. 49: Classification error rate for three implemented classifiers with different dimension reduction techniques [86].....	57
Figure 2. 50: Classification error for four classifiers under the same force [87]	58
Figure 2. 51: Classification error for two classifiers with different features set [89]	59
Figure 2. 52: Pattern Recognition Architecture	59
Figure 2. 53: Data Segmentation via windowing (a) disjoint windowing	60
Figure 2. 54: Three leading TFD features (a) STFT (b) WT (c) WPT [95].....	65
Figure 2. 55: The architecture of primary extreme learning machine network.....	72
Figure 2. 56: Support vector machine classifier [110].....	77
Figure 3. 1: Four key steps of feature selection [131]	86
Figure 3. 2: Ten targeted finger movements [75]	90
Figure 3. 3: CDF values versus P values	91
Figure 3. 4: Misclassification Error	92
Figure 3. 5: Forward Sequential Feature Selection.....	93
Figure 3. 6: MCE for resubstitution and cross folded (QDA).....	93
Figure 3. 7: MCE for resubstitution and cross-folded (ELM)	94
Figure 3. 8: The structure of natural brain [196].....	105
Figure 3. 9: The structure of neural networks	105
Figure 3. 10: The single layer feedforward networks architecture [206].....	107
Figure 3. 11: The multi-layer fully connected feedforward networks architecture [206].....	108
Figure 3. 12: The multi-layer not fully connected feedforward networks architecture [206].....	108
Figure 3. 13: The recurrent network's architecture [206].....	109
Figure 3. 14: The lattice networks architecture [206]	109
Figure 4. 1: Procedures for sparse autoencoder signal representation	131
Figure 4. 2: Procedures of superimposed signal representations suggested model	132
Figure 4. 3: Placement of electrodes [75]	135
Figure 4. 4: Targeted ten different finger motions [75].....	135
Figure 4. 5 : ANOVA values for different Classifiers for individual signal representation	138
Figure 4. 6: Confidence Intervals for used classifiers for individual signal representation.....	139
Figure 4. 7: Electromyography signal for ten-finger movements	140
Figure 4. 8: Testing accuracy versus window size for superimposed signal representation.....	141
Figure 4. 9: Architecture of pattern recognition with AW-ELM	142
Figure 5. 1: Procedures of the suggested PCA feature learning model	149
Figure 5. 2: Procedures of the proposed ICA feature learning model	149
Figure 5. 3: Number of learnt features for each data representation.....	153
Figure 5. 4: Error rate versus number of hidden nodes for offline classification [81].....	156
Figure 5. 5: Average accuracy versus ten classes for RBF-ELM & OS-ELM [81].....	156

Figure 6. 1: Procedures of suggested SOM feature learning model	161
Figure 6. 2: Mapping input space to output space in SOM.....	162
Figure 6. 3: Pattern Recognition Model.....	171
Figure 6. 4: Confidence interval for four classifiers	173
Figure 7. 1: Procedures for data augmented tensor representation deep learning model.....	179
Figure 7. 2: Electrodes allocation for amputee participants [87]	183
Figure 7. 3: Comparison between suggested deep learning model and pattern recognition in ten finger movement.....	186
Figure 7. 4: Comparison between suggested deep learning model and pattern recognition in six-finger movement.....	186
Figure 7. 5: Average accuracy grouped by the number of classes for various dimension reduction [80] ..	188
Figure 7. 6: Average accuracy grouped by five classes for different dimension reduction algorithms [80]	189

List of Tables

Table 1. 1: International Disability Occurrence for Adults.....	2
Table 1. 2: The Washington Group study results [6].....	2
Table 1. 3: The disability percentages for children [8].....	3
Table 4. 1: Autoencoder testing and training accuracy for individual signal representation.....	137
Table 4. 2: Autoencoder classifier fusion testing and training accuracy for individual signal representation	139
Table 4. 3: Training and testing accuracies for three-fold cross validations for superimposed signal representation.....	140
Table 4. 4: Training and testing accuracies for ten-fold cross validations for superimposed signal representation.....	140
Table 4. 5: Average testing accuracy for three activation functions.....	143
Table 4. 6: Average processing time for ELM with different activation functions.....	143
Table 4. 7: P-value for AW-ELM and other activation functions.....	143
Table 5. 1: PCA training and testing accuracy.....	151
Table 5. 2: Classifier fusion training and testing accuracy for PCA.....	152
Table 5. 3: Feature indexing training and testing accuracy.....	153
Table 5. 4: ICA training and testing accuracy.....	154
Table 5. 5: Classifier fusion training and testing accuracy for ICA.....	154
Table 6. 1: SOM deep learning training and testing accuracy.....	170
Table 6. 2: Confidence interval for various classifiers.....	172
Table 6. 3: Classifier fusion training and testing accuracy.....	173
Table 6. 4: Average accuracy and simulation time for SRDA and RLDA.....	174
Table 7. 1: Training and Testing accuracies for all implemented datasets.....	185
Table 7. 2: Different classes' combinations were implicated in the study.....	188

List of Abbreviations

ANN	Artificial neural network
ANOVA	Analysis of variance
AR	Autoregressive
AW-ELM	Adaptive wavelet extreme learning machine
CNN	Convolutional neural network
CDF	Cumulative distribution function
CMC	Carpometacarpal
CPD	Canonical polyadic decomposition
CTC	Connectionist temporal classification
DA	Discriminant analysis
DCS	Dynamic classifier selection
DCS-LA	Dynamic classifier selection local accuracy
DFT	Discrete Fourier transform
DTNN	Deep tensor neural network
ECG	Electrocardiogram
ECOG	Electrocorticographic
ECRB	Extensor carpi radialis brevis
ECRL	Extensor carpi radialis longus
ECU	Extensor carpi ulnaris
EDA	Electrodermal activity
EEG	Electroencephalogram
ELM	Extreme learning machine
EMG	Electromyography
EOG	Electrooculography
FCU	Flexor carpi ulnaris
FFT	Fast Fourier transform

FLDA	Fuzzy LDA
FLN	Functional-link net
FNN	Feed-forward neural networks
FNPA	Fuzzy neighbourhood preserving analysis
FP	Feature projection
FS	Feature selection
FT	Fourier transform
GSR	Galvanic skin response
HME	Hierarchical mixture of experts
HMM	Hidden Markov model
HTD	Hjorth time domain
ICA	Independent component analysis
KNN	k nearest neighbourhood
K-SVD	K-singular value decomposition
LDA	Linear discriminant analysis
LIBSVM	Library support vector machine
LinELM	Linear extreme learning machine
LLE	Local linear embedding
LPM	Linear programming machine
LPP	Locality preserving projection
LS-SVM	Least-square SVM
LSTM	Long Short-Term Memory
MAV	Mean absolute value
MAVs	Mean absolute value slope
MC	Markov chain
MCA	Mutual components analysis
MCE	Misclassification error
MCMC	Markov chain Monte Carlo

MDF	Median frequency
MDP	Markov decision process
MEG	Magnetoencephalogram
MES	Myoelectric signal
MI	Mutual information
MLP	Multilayer perceptron
MMG	Mechanomyogram
MNF	Mean frequency
NB	Naive bayes
NMF	Nonnegative matrix Factorization
OFNDA	Orthogonal fuzzy neighbourhood discriminant analysis
OLDA	Orthogonal LDA
OS-ELM	Online sequential extreme learning machine
PCA	Principal component analysis
PNN	Probabilistic neural network
Poly-ELM	Polynomial extreme learning machine
PSO	Particle swarm optimization
QDA	Quadrature discriminant analysis
QPC	Quadratic phase coupling
RBF	Restricted Boltzmann machine
RBN	Radial basis network
RegTree	Regression Tree
ReLU	Rectified linear unit
RF	Random forest
RGB	Red Green Blue
RLDA	Regularised linear discriminant analysis
RMS	Root mean square

RVFLN	Random vector functional-link net
SA-EELM	Self-adaptive evolutionally extreme learning machine
SA-SVM	Self-advising SVM
Sig-ELM	Sigmoid extreme learning machine
SFLNs	Single-hidden-layer feed-forward networks
SL	Softmax layer
SOM	Self-organising map
SPCA	Sparse PCA
SR	Spectral regression
SS	Sample skewness
SSC	Slope sign changes
SRDA	Spectral regression discriminant analysis
STFT	Short-time Fourier transform
SVD	Singular value decomposition
SVM	Support vector machine
TD	Time domain
TDD	Time-domain descriptors
TDNN	Time delay neural network
ULDA	Uncorrelated linear discriminant analysis
USELM	Unsupervised extreme learning machine
WAMP	Willison amplitude
W-ELM	Wavelet extreme learning machine
WHO	World health organisation
WL	Waveform length
WPT	Wavelet packet transform
WT	Wavelet transform
ZC	Zero crossing

This page is intentionally left blank

Abstract

A wide gap exists between clinical application results and those from laboratory observations concerning hand rehabilitation devices. In most instances, laboratory observations show superior outcomes the real-time applications demonstrate poor consequences. The robust nature of the electromyography signal and limited laboratory applications are the principal reasons for the gap. This thesis aims to introduce and develop a deep learning model that is capable of learning features from biosignals.

The deep learning model is expected to tame the variable nature of the electromyography signal which will lead to the best available outcomes. Furthermore, the suggested deep learning scheme will be trained to be skilled in learning the best features that match the biosignal application regardless of the number of classes. Moreover, traditional feature extraction is time consuming and extremely reliant on the user's experience and the application. The objective of this research is accomplished via the following four implemented models.

1. Developing a deep learning model via implementing a two-stage autoencoder along with applying different signal representations like spectrogram, wavelet and wavelet packet to tame variations of the electromyography signal. Support vector machine, extreme learning machine with two activation functions (sigmoid and radial basis function) and softmax layer were used for classifications. Moreover, the classifier fusion layer achieved testing accuracy of more than 92% and training attained more than 98%. The same dataset was implemented for superimposed signal representations for two stages autoencoder and softmax layer, support vector machine, k-nearest neighbor and discriminant analysis for classification besides the classifier fusion which led to testing accuracy of more than 90%.
2. Presenting principal component analysis and independent component analysis for feature learning purposes after applying different signal representations algorithms such as spectrogram, wavelet and wavelet packet. Discriminant analysis, extreme learning machine and support vector machine were used for classification. Furthermore, the two proposed models showed acceptable accuracy along

with shorter simulation time. The testing accuracy achieved more than 90% by implementing a classifier fusion layer. Manhattan index was estimated for all features and only the top 50 Manhattan index features were included to decrease the simulation time while attaining acceptable accuracy values.

3. Introducing a self-organising map for deep learning whereby the biosignal was represented by spectrograms, wavelet and wavelet packet. The presented biosignal was introduced to a layer of self-organising map then the suggested system performance was evaluated by extreme learning machine, self-adaptive evolutionally extreme learning machine, discriminant analysis and support vector machine for classification. Adding a classifier fusion layer increased the testing accuracy to 96.60% for ten-finger movements and 99.73% for training. The proposed system showed superior behavior regarding accuracy and simulation time.
4. Presenting a deep learning model where 1) the data was augmented after representing the biosignal by a spectrogram, 2) the augmented signal was represented by a tensor, and finally 3) The signal was introduced to the two-stage autoencoder. The same dataset was used with traditional pattern recognition for comparison purposes. Classifier fusion layer was executed in deep learning scheme whereby the ten-finger movements achieved 90.25% and 87.11% attained by pattern recognition. Besides, the six finger movement dataset was acquired from amputee participants and accomplished 91.85% for deep learning and reached 89.64% for traditional pattern recognition. Furthermore, different datasets for different applications were tested using the recommended deep learning model. Eventually, feeding the deep learning model with various datasets for different applications afforded the model with higher fidelity, combined with real outcomes and generalization.