
Evolutionary Wavelet Neural Networks in Data Classification and Dynamic Control

Maryam Mahsal KHAN

BSc(CISE)(Pakistan); MSc(EE)(Malaysia)

*A thesis submitted in partial fulfilment of the requirements
for the degree of Doctor of Philosophy in Computer Science*

February 23, 2018

This research was supported by an Australian Government Research Training
Program (RTP) Scholarship

Declaration of Authorship

Statement of Originality

I, Maryam Mahsal KHAN, hereby certify that the work embodied in the thesis is my own work, conducted under normal supervision.

The thesis contains no material which has been accepted, or is being examined, for the award of any other degree or diploma in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. I give consent to the final version of my thesis being made available worldwide when deposited in the University's Digital Repository, subject to the provisions of the Copyright Act 1968 and any approved embargo.

Signed:

Date:

Acknowledgements

First praise to Allah (God), the Almighty, for giving me this opportunity and for the strength and patience required in my quest for knowledge.

I would like to express my deepest gratitude to my supervisors, Stephan K. Chalup and Alexandre Mendes. Their invaluable guidance, insightful suggestions, experience and support have played a critical role in my research endeavour. Special thanks to the examiners for their comments and feedback.

My appreciation to colleagues in the Interdisciplinary Machine Learning Research Group for their effective discussion and ideas and many thanks to Aaron Scott for his support on the HPC grid. I would like to thank Ping Zhang for sharing the breast cancer dataset for my research and to all the paper reviewers who have provided outstanding feedback. Many thanks to the Australian Government Research Training Program for not only providing funding to undertake this research but also for giving me the opportunity to attend conferences and workshops.

'All work and no play makes Jill a dull girl'. Thanks to the different clubs at the University of Newcastle, especially NUSA and NUPSA, for organising various events, and to the Student Association of Pakistan (SAP) for a much-needed form of escape from my studies. Many thanks to my friends at Newcastle for making life more exciting through many get-togethers. Batool Naveed's family deserves a special thanks because they have been like a second family to me. Indeed, you all were better than any energy booster.

Finally, I would like to deeply acknowledge my dear parents, my father Mahsal Khan and my mother Shagufta Nasreen, who taught me the value of education and hard work. Where I stand today is due to their selfless efforts, prayer and encouragement. I hope to be as dedicated as they are. And to my loving husband Naseer Ahmed Khan for his ongoing support, and to my sweet little naughty daughter Rushda for adding colour to my life.

Thank you.

The University of Newcastle

Abstract

Faculty of Built Environment and Engineering

School of Electrical Engineering & Computing

Doctor of Philosophy in Computer Science

Evolutionary Wavelet Neural Networks in Data Classification and Dynamic Control

by Maryam Mahsal KHAN

A wavelet neural network (WNN) is a combination of a neural network with wavelet functions, and belongs to a special class of neural networks in the field of machine learning. The interesting aspect of this type of network is that it has a single hidden layer. The advantages of such an inherent property are two-fold: fast convergence speeds and easy assessment of each neuron's contribution towards prediction.

Despite the above advantages of WNNs, the optimisation of their parameters and the estimation of the number of hidden neurons have significant effects on their performance. Not all WNN parameters are easily differentiable and are therefore usually excluded from the training process. Currently, standard gradient-based algorithms are used to optimise the different parameters of networks. Moreover, the initialisation of hidden neurons plays a critical role in capturing the variability of data.

Evolutionary algorithms have been used as a gradient-free optimisation method in many research problems where differentiability is unavailable or derivatives are unreliable or impractical to obtain. Thus, evolutionary algorithms was the effective choice for WNN parameter optimisation. Furthermore, in order to devise a bloat-free evolutionary programming method, a Cartesian genetic programming (CGP) model was used. Such models are based on the concept of using and evolving fixed resources

such as nodes and their connections links. This phenomenon proves beneficial where adaptability of hidden neurons is required, as its quantification varies from one system to another.

The proposed evolutionary WNN (EWNN) was first applied to a standard two-spiral task. This benchmark task provided a clear understanding of the operation and response of EWNNs, which highlights their potential for separating non-linear classes.

An EWNN was then applied to the classification of three publicly-available biomedical datasets on breast cancer and Parkinson's disease. The process of feature pruning during the evolutionary process, and the effects of training all of the network parameters, were studied in detail.

To further improve the classification performance of EWNNs, an ensemble EWNN (EWNN-e) was proposed. In this method, a genetic algorithm was used to prune trained EWNN classifiers for the three previously-investigated datasets. The EWNN-e was found to be even more accurate than the independent EWNN classifier.

The performance of EWNNs in learning patterns of control behaviour in a benchmark control problem —the acrobot— was the final focus of this thesis. The performance of any reinforcement learning algorithm is dependent on the space domain it operates in, i.e. discrete or continuous, whereby a discrete action space is significantly less challenging than a continuous one. In the context of EWNNs, both discrete and continuous action spaces were investigated. The performance of EWNNs were compared with the state-of-the-art deep RL algorithm. The EWNNs produced robust acrobot controllers that were independent of the type of action space domain.

Keywords: Wavelet neural networks, Cartesian genetic programming, prediction models, intelligent controllers.

Contents

| | |
|---|------------|
| Declaration of Authorship | iii |
| Acknowledgements | v |
| Abstract | vii |
| List of Figures | xv |
| List of Tables | xix |
| 1 Introduction | 1 |
| 1.1 History of Wavelets | 2 |
| 1.2 Major Challenges | 3 |
| 1.3 Research Objectives | 5 |
| 1.4 Publications and Achievements | 7 |
| 1.5 Thesis Overview and Summary | 10 |
| 1.6 Chapter Summary | 11 |
| 2 Technical Background and Literature Review | 13 |
| 2.1 Wavelet Neural Networks | 13 |
| 2.1.1 Properties of WNNs | 14 |
| 2.1.2 Learning in WNNs | 17 |
| 2.1.3 Initialisation of the WNN | 18 |
| 2.1.4 Applications of WNN | 20 |
| 2.2 Neuroevolutionary Algorithms | 21 |
| 2.2.1 Direct representations | 21 |

| | | |
|----------|---|-----------|
| 2.2.2 | Developmental Representation | 24 |
| 2.2.3 | Indirect or Implicit Representation | 25 |
| 2.3 | Cartesian Genetic Programming | 25 |
| 2.4 | Chapter Summary | 27 |
| 3 | Evolutionary Wavelet Neural Network | 29 |
| 3.1 | Insight to WNNs | 29 |
| 3.2 | EWNN structure and evolutionary operators | 35 |
| 3.2.1 | Algorithm Settings | 35 |
| 3.2.1.1 | Common Genome Structure | 35 |
| 3.2.1.2 | Evolution Strategy | 36 |
| 3.2.1.3 | Initialisation of the network | 37 |
| 3.2.1.4 | Mutation | 37 |
| 3.3 | Proposed Algorithms | 39 |
| 3.3.1 | Algorithm A1: One-dimensional WNN | 39 |
| 3.3.2 | Algorithm A2: Multi-dimensional WNN | 41 |
| 3.3.3 | Algorithm A3: Multi-dimensional WNN with Rotation | 43 |
| 3.3.4 | Algorithm A4: Self-evolving Multi-dimensional WNN with Ro- tation | 45 |
| 3.3.5 | Derived Algorithms | 46 |
| 3.4 | Two-Spiral Task | 51 |
| 3.4.1 | Configuration 1: Constructing a WNN using (α, β) with a stan- dard ψ (Algorithm A2) | 53 |
| 3.4.2 | Configuration 2: Constructing a WNN using a radial ψ (Algo- rithm A2) | 53 |
| 3.4.3 | Configuration 3: Constructing a WNN using (α, β, R) with a Standard ψ (Algorithm A3) | 55 |
| 3.5 | Chapter Summary | 56 |

| | | |
|----------|---|-----------|
| 4 | Evolutionary One-dimensional Wavelet Neural Networks | 57 |
| 4.1 | Breast Cancer Detection | 57 |
| 4.1.1 | Database and features utilised | 58 |
| 4.1.2 | Related Work | 60 |
| 4.1.3 | Training and testing sets | 62 |
| 4.1.3.1 | Training on 70% of the data | 62 |
| 4.1.3.2 | 10-fold cross-validation | 62 |
| 4.2 | Experimental setup | 63 |
| 4.2.1 | CGPANN parameters | 63 |
| 4.2.2 | EWNN parameters | 64 |
| 4.2.3 | NEAT parameters | 65 |
| 4.3 | Results and Discussion | 66 |
| 4.4 | Chapter Summary | 69 |
| 5 | Evolutionary Multi-dimensional Wavelet Neural Networks | 71 |
| 5.1 | Breast Cancer Classification | 71 |
| 5.1.1 | Training and Testing Sets | 71 |
| 5.1.2 | Experimental Setup | 72 |
| 5.1.3 | Results and Discussion | 72 |
| 5.2 | Sakar's Parkinson's Disease Dataset Classification | 76 |
| 5.2.1 | Database and Features | 77 |
| 5.2.2 | Training and Testing Sets | 79 |
| 5.2.3 | Experimental Setup | 80 |
| 5.2.4 | Results and Discussion | 81 |
| 5.3 | Little's Parkinson's Disease Dataset Classification | 83 |
| 5.3.1 | Related Work | 85 |
| 5.3.2 | Experimental Setup | 87 |
| 5.3.3 | Results and Discussion | 87 |
| 5.4 | Chapter Summary | 88 |

| | | |
|----------|--|------------|
| 6 | Evolutionary Wavelet Neural Network Ensembles | 91 |
| 6.1 | Ensemble Pruning | 92 |
| 6.2 | Experimental Setup | 93 |
| 6.2.1 | Datasets | 94 |
| 6.2.2 | Parameter Settings | 95 |
| 6.3 | Results and Discussion | 96 |
| 6.4 | Chapter Summary | 101 |
| 7 | Evolutionary Wavelet Neural Networks in Control | 103 |
| 7.1 | Background | 103 |
| 7.1.1 | Deep Reinforcement Learning | 103 |
| 7.1.2 | Neuroevolution in Control | 104 |
| 7.2 | The Acrobot Control Problem | 105 |
| 7.2.1 | Related Work | 107 |
| 7.3 | Experimental Setup | 108 |
| 7.3.1 | EWNN Configuration | 108 |
| 7.3.2 | DNN Configuration | 110 |
| 7.4 | Results & Discussion | 112 |
| 7.4.1 | EWNN - Discrete action space | 112 |
| 7.4.2 | EWNN - Continuous action space | 116 |
| 7.4.3 | DNN - TRPO RL algorithm | 118 |
| 7.5 | Chapter Summary | 120 |
| 8 | Conclusion | 123 |
| 8.1 | Main Research Contributions | 125 |
| 8.2 | Further Work | 128 |
| 8.2.1 | Reducing Rotation Computational Time | 128 |
| 8.2.2 | EWNNs on Huge Datasets | 128 |
| 8.2.3 | Other Ensemble Approaches | 129 |
| 8.2.4 | Analysis of Complex Control Tasks | 129 |

| | |
|--|-----|
| 8.2.5 Other Research Avenues | 130 |
|--|-----|

| | |
|---------------------|------------|
| Bibliography | 131 |
|---------------------|------------|

List of Figures

| | | |
|------|---|----|
| 1.1 | Different types of wavelet functions and their frequency responses. . . . | 2 |
| 1.2 | Structure of a Wavelet neural network. | 4 |
| 2.1 | Properties of a Wavelet neural network. | 14 |
| 2.2 | Different types of wavelons. | 16 |
| 2.3 | Different responses of a Morelet wavelet function. | 16 |
| 2.4 | Example of a CGP genotype and phenotype. | 26 |
| 3.1 | Effect of changing the dilation parameter in a WNN. | 30 |
| 3.2 | Effect of changing the translation parameter in a WNN. | 31 |
| 3.3 | Effect of changing the rotation parameter in a WNN. | 31 |
| 3.4 | Three different types of wavelet functions and their frequency responses. | 32 |
| 3.5 | Example of a one-dimensional and multi-dimensional dilations and translations. | 33 |
| 3.6 | Radial responses of three common wavelet functions. | 34 |
| 3.7 | EWNN genome specification. | 35 |
| 3.8 | Standard representation of wavelons. | 39 |
| 3.9 | General composition of a one-dimensional EWNN wavelon genome. | 40 |
| 3.10 | General composition of a one-dimensional EWNN output genome. | 41 |
| 3.11 | Example of a one-dimensional EWNN genome. | 42 |
| 3.12 | General composition of a multi-dimensional EWNN wavelon genome. | 43 |
| 3.13 | Example of a multi-dimensional EWNN genome. | 44 |
| 3.14 | General composition of a multi-dimensional EWNN genome with rotation parameter. | 45 |

| | | |
|------|--|-----|
| 3.15 | General composition of a self-evolving EWNN output genome. | 46 |
| 3.16 | Population of a self-evolving EWNN structures with random outputs <i>OL</i> | 47 |
| 3.17 | Training set of points of the two-spiral benchmark task. | 52 |
| 3.18 | Responses of a EWNN with 20 wavelons in the hidden layer on the two-spiral task with a Morelet wavelet function. | 54 |
| 3.19 | Responses of a EWNN with 30 wavelons in the hidden layer on the two-spiral task with double derivative of Gaussian as a wavelet function. | 54 |
| 3.20 | Responses of a two-wavelon EWNN on the two-spiral task with radi- ally symmetric Morelet wavelet functions. | 55 |
| 3.21 | Responses of a EWNN with 20 wavelons in the hidden layer on the two-spiral task with rotation parameter and a Morelet wavelet function. | 56 |
| 4.1 | Feature selection of the evolved one-dimensional EWNNs. | 67 |
| 5.1 | Fitness graph of two structures of EWNN on the DDSM dataset with rotation enabled. | 74 |
| 5.2 | Comparison of the average number of mutations of EWNN parameters on the DDSM dataset. | 75 |
| 5.3 | Comparison of the frequency of the wavelet function ψ used in the final genotypes of the winning structures on the DDSM dataset. | 76 |
| 5.4 | EWNN response (rotation enabled) of the best genotypes for both train- ing strategies on the Sakar's Parkinson's dataset. | 82 |
| 6.1 | Cumulative number of occurrences of the different dimensions of wavelons for stand-alone EWNN; and for the ensemble EWNN-e, for the three datasets using 10-fold cross-validation. | 98 |
| 6.2 | Average number of connections per feature within the EWNN and its ensemble EWNN-e, for DDSM, LPD, and SPD datasets, respectively. . . | 99 |
| 7.1 | The acrobat control problem. | 105 |
| 7.2 | Torque responses of acrobat controllers with different wavelet functions. | 113 |

| | | |
|------|--|-----|
| 7.3 | Average swing-up times at each evolutionary step of fifty independent evolutionary runs. | 114 |
| 7.4 | Behaviour of the best EWNN swing-up controller in the height task and hand-stand acrobot control tasks. | 115 |
| 7.5 | Average minimum and maximum torque ranges of the EWNNs for the different wavelet functions. | 117 |
| 7.6 | Average swing-up times at each evolutionary step for <i>height</i> and <i>hand-stand</i> acrobot control tasks. | 118 |
| 7.7 | Behaviour of the best EWNN swing-up controller in the <i>height</i> and <i>hand-stand</i> acrobot control tasks. | 119 |
| 7.8 | Average swing-up times and best swing-up times on the <i>height</i> task for a deep neural network. | 120 |
| 7.9 | Average swing-up times and best swing-up times for a deep neural network on the acrobot <i>hand-stand</i> task. | 121 |
| 7.10 | Snapshots of a successful acrobot DNN controller on the <i>height</i> task. . . | 121 |

List of Tables

| | | |
|-----|---|----|
| 4.1 | Features and description used in the Digital Database for Screening Mammography dataset. | 59 |
| 4.2 | Performance of CGPANN with a 70%/30% split between training and testing samples where the best was tested using 10-fold cross-validation. | 63 |
| 4.3 | Performance of EWNN & EWNN with linearity disabled on 70% Training and 30% Testing Dataset where the best is tested using 10-fold cross-validation. | 65 |
| 4.4 | NEAT algorithm parameters | 66 |
| 4.5 | Performance of NEAT using 70/30 split and 10-fold cross-validation strategies on the DDSM dataset. | 66 |
| 4.6 | Performance of neuroevolutionary algorithms. | 68 |
| 4.7 | Comparison of different classifiers on the DDSM using 6 features | 70 |
| 5.1 | Performance of a multi-dimensional EWNN on the DDSM dataset using 50/50 split and 10-fold cross-validation strategies. | 73 |
| 5.2 | Performance comparison of different classifiers on the DDSM. | 77 |
| 5.3 | Parkinson's disease dataset features and their descriptions. | 78 |
| 5.4 | Performance of multi-dimensional EWNN on the Sakar's Parkinson's dataset using 10-fold cross-validation and leave-one-subject-out strategies. | 81 |
| 5.5 | Performance of SVM classifier on leave-one-subject-out (LOSO) and 10-fold cross-validation test strategies with Sakar's Parkinson's dataset. | 81 |

| | | |
|-----|--|-----|
| 5.6 | Comparison of techniques using leave-one-subject-out and 10-fold cross-validation strategies with Sakar's Parkinson's Disease classification. . . . | 83 |
| 5.7 | Parkinson's dataset features and their descriptions. | 84 |
| 5.8 | Performance of multi-dimensional EWNN on Little's Parkinson's dataset, using a 10-fold cross-validation strategy. | 88 |
| 5.9 | Comparison of techniques using 10-fold cross-validation strategies in Little's Parkinson's disease classification. | 89 |
| 6.1 | Dataset characteristics. | 94 |
| 6.2 | Main parameter settings of the evolutionary wavelet neural networks for the different case studies. | 96 |
| 6.3 | Performance of the ensemble EWNN on the different case studies. | 97 |
| 6.4 | Average number of active EWNNs in an ensemble on 10-fold cross-validated datasets. | 100 |
| 6.5 | Comparison of different ensemble and clustering classifiers on the DDSM, LPD and SPD datasets. | 100 |
| 7.1 | Parameters for an acrobot control problem. | 106 |
| 7.2 | Performance of EWNN on acrobot swing-up for the <i>height</i> and <i>hand-stand</i> tasks in a discrete action space, under different parameter configurations of torque range, type of wavelet functions and maximum numbers of wavelons. | 111 |
| 7.3 | Performance of EWNN on acrobot swing-up for the <i>height</i> and <i>hand-stand</i> tasks in continuous action space, under different type of wavelet functions and maximum numbers of wavelons. | 112 |
| 7.4 | Parameter settings for Trust Region Policy Optimization RL algorithm on the acrobot control tasks. | 112 |
| 7.5 | Comparison of acrobot swing-up times by several methods reported in the literature. | 122 |

List of Abbreviations

| | |
|---------------|--|
| ANN | Artificial Neural Network |
| CGP | Cartesian Genetic Programming |
| CGPANN | Cartesian Genetic Programming Artificial Neural Network |
| DDSM | Digital Database of Screening Mammography |
| DNN | Deep Neural Network |
| EWNN | Evolutionary Wavelet Neural Network |
| EWNN-e | Evolutionary Wavelet Neural Network Ensembles |
| LCA | Latent Class Analysis |
| NEAT | Neuro Evolution of Augmented Topology |
| NN | Neural Network |
| PD | Parkinson's Disease |
| RL | Reinforcement Learning |
| SVM | Support Vector Machine |
| TRPO | Trust Region Policy Optimization |
| WNN | Wavelet Neural Network |
| WT | Wavelet Transform |

