

2006

A new class of convolutional neural networks based on shunting inhibition with applications to visual pattern recognition

Fok Hing Chi Tivive
University of Wollongong, tivive@uow.edu.au

Follow this and additional works at: <https://ro.uow.edu.au/theses>

University of Wollongong

Copyright Warning

You may print or download ONE copy of this document for the purpose of your own research or study. The University does not authorise you to copy, communicate or otherwise make available electronically to any other person any copyright material contained on this site.

You are reminded of the following: This work is copyright. Apart from any use permitted under the Copyright Act 1968, no part of this work may be reproduced by any process, nor may any other exclusive right be exercised, without the permission of the author. Copyright owners are entitled to take legal action against persons who infringe their copyright. A reproduction of material that is protected by copyright may be a copyright infringement. A court may impose penalties and award damages in relation to offences and infringements relating to copyright material.

Higher penalties may apply, and higher damages may be awarded, for offences and infringements involving the conversion of material into digital or electronic form.

Unless otherwise indicated, the views expressed in this thesis are those of the author and do not necessarily represent the views of the University of Wollongong.

Recommended Citation

Tivive, Fok Hing Chi, A new class of convolutional neural networks based on shunting inhibition with applications to visual pattern recognition, PhD thesis, School of Electrical, Computer and Telecommunications Engineering, University of Wollongong, 2006. <http://ro.uow.edu.au/theses/540>

Research Online is the open access institutional repository for the University of Wollongong. For further information contact the UOW Library: research-pubs@uow.edu.au

NOTE

This online version of the thesis may have different page formatting and pagination from the paper copy held in the University of Wollongong Library.

UNIVERSITY OF WOLLONGONG

COPYRIGHT WARNING

You may print or download ONE copy of this document for the purpose of your own research or study. The University does not authorise you to copy, communicate or otherwise make available electronically to any other person any copyright material contained on this site. You are reminded of the following:

Copyright owners are entitled to take legal action against persons who infringe their copyright. A reproduction of material that is protected by copyright may be a copyright infringement. A court may impose penalties and award damages in relation to offences and infringements relating to copyright material. Higher penalties may apply, and higher damages may be awarded, for offences and infringements involving the conversion of material into digital or electronic form.

**A NEW CLASS OF CONVOLUTIONAL NEURAL NETWORKS
BASED ON SHUNTING INHIBITION WITH APPLICATIONS TO
VISUAL PATTERN RECOGNITION**

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

DOCTOR OF PHILOSOPHY

from

UNIVERSITY OF WOLLONGONG

by

FOK HING CHI TIVIVE

B. Eng. (Hons.)

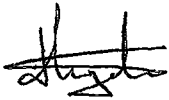
Supervisor : Prof. Abdessalam Bouzerdoum

School of Electrical, Computer and Telecommunications Engineering

March 2006

Certification

I, Fok Hing Chi Tivive, declare that this thesis, submitted in fulfilment of the requirements for the award of Doctor of Philosophy, in the School of Electrical, Computer and Telecommunications Engineering, University of Wollongong, is wholly my own work unless otherwise referenced or acknowledged. The document has not been submitted for qualifications at any other academic institution.



Fok Hing Chi Tivive

March 2006

Abstract

In the contemporary era of increased information overload, there is a growing interest in a new class of computational intelligence systems. These systems have been proven as powerful and versatile computational tools for solving certain types of problems that are too complex to be analyzed with traditional analytical means. Inspired by the computational mechanism of the human brain, many researchers have looked into neurobiology for new inspiration to solve more complex problems than those based on traditional computational techniques. Artificial neural networks, evolving from neuro-biological insights, give computer systems an amazing capability to actually learn from input data to generate solutions for problems that are too abstract to be understood or too resource-intensive to tackle. Although neural networks have been applied with success in many industries, there is a continuing demand for new types of hierarchical artificial neural networks that can overcome some of the drawbacks of the earlier models.

This thesis presents a new class of convolutional neural networks based on the physiologically plausible mechanism of shunting inhibition with its various systematic connection schemes. The network has a generic architecture in which shunting inhibitory neurons are used as feature extraction elements. A series of training algorithms, ranging from first-order gradient methods to Quasi-Newton and hybrid methods, have been implemented to adapt the synaptic weights of the developed networks; all of them have been successfully used to train the convolutional neural networks for a classification task.

To demonstrate their capability in real life applications, the convolutional networks are employed in a face detection system and a handwritten digit recognition system. The face detector has 383 trainable network parameters and achieves a detection rate of 98% for detecting human faces on a large set of unconstrained and complex images. The handwritten digit recognition system, on the other hand, has 2722 trainable parameters, and its classification rate is 97.3% for recognizing human

handwritten numerals. Besides these two applications, the developed network is analyzed for its built-in invariance, and it is implemented as a rotation invariant face classifier. The network achieves a classification rate of 97.3% in the rotation range $\pm 90^\circ$, and for 360° in-plane rotation, it has a correct detection rate of 93.6% at 5% false detection rate. These classification results demonstrate that the new class of convolutional neural networks has excellent generalization capability and achieves rotation invariance by adapting its connection weight matrices (receptive fields) as invariant feature detectors.

Acknowledgements

This thesis could not be written without the help and support of the following people. First, I would like to express my deepest appreciation and gratitude to my supervisor, Prof. Abdessalam Bouzerdoun, for his advice and permanent guidance throughout my research. I was fortunate to have him as my supervisor for both undergraduate and postgraduate studies and have his support during all these years. I would like to thank my best friend, Dr. Son Lam Phung, for his expert advice in Matlab, his help in image processing and his time for proof-reading this thesis. A special thank goes to Dr. Farid Boussaid who has shared his past experience of how to manage the time for writing a Ph.D thesis and his motivating phone calls. I would like to thank my girlfriend Bonnie for her love and daily phone call and emails which have motivated me to complete my Ph.D. And finally, I would like to dedicate this thesis to my parents, brothers and my closest aunty Lee for their support and encouragements.

Publications

- F. H. C. Tivive and A. Bouzerdoun, “Handwritten digit recognition based on shunting inhibitory convolutional neural networks,” in *Proc. of the Workshop on Learning Algorithms for Pattern Recognition in conjunction with the 18th Australian Joint Conference on Artificial Intelligence*, 2005, pp. 72-77.
- F. H. C. Tivive and A. Bouzerdoun, “A fast neural-based eye detection system,” in *Proc. of the International Symposium on Intelligent Signal Processing and Communications Systems* , 2005, pp. 641-644.
- F. H. C. Tivive and A. Bouzerdoun, “An eye feature detector based on convolutional neural network,” in *Proc. of the Eighth International Symposium on Signal Processing and its Applications*, 2005, pp. 90-93.
- F. H. C. Tivive and A. Bouzerdoun, “Efficient training algorithms for a class of shunting inhibitory convolutional neural networks,” *IEEE Transactions on Neural Networks*, vol. 16, no. 3, pp. 541-556, 2005.
- F. H. C. Tivive and A. Bouzerdoun, “A face detection system using shunting inhibitory convolutional neural network,” in *Proc. of the International Joint Conference on Neural Networks*, vol. 4, 2004, pp. 2571-2575.
- F. H. C. Tivive and A. Bouzerdoun, “A new class of convolutional neural networks (siconnets) and their application to face detection,” in *Proc. of the International Joint Conference on Neural Networks*, vol. 3, 2003, pp. 2157-2162.

Contents

Abstract	iii
Acknowledgements	v
Publications	vi
Nomenclature	xxiii
1 Introduction	1
1.1 Motivation and Significance	2
1.2 Research Objectives	3
1.3 Outline of Thesis	4
2 A Review of Artificial Neural Networks	6
2.1 Introduction	6
2.2 Artificial Neural Networks	7
2.2.1 From Biological to Artificial Neuron	8
2.2.2 Network Topologies	11
2.2.3 Multilayer Perceptrons	11
2.3 Convolutional Neural Networks	14
2.3.1 Neocognitron	15
2.3.2 LeNet Network	21
2.4 Applications of Convolutional Neural Networks	25
2.4.1 Visual Documents Analysis	25
2.4.2 Facial Expression Analysis	27
2.4.3 Biomedical Image Processing	28
2.4.4 Automatic Target Recognition	30
2.4.5 Biometric Recognition	31

2.4.6	Neural Based Tracking Systems	34
2.4.7	Other Systems	35
2.5	Conclusion	36
3	A New Class of Convolutional Neural Networks	37
3.1	Introduction	37
3.2	Description of the Network Architecture	38
3.2.1	Sub-sampling Process	40
3.2.2	Output Layer of SICoNNet	42
3.3	Connection Strategies	43
3.4	Shunting Inhibitory Neuron Model	46
3.4.1	Lateral and Shunting Inhibition Mechanisms	47
3.4.2	Decision Surfaces of the SIANN	50
3.5	Structural Difference Between the New Architecture and the Existing CoNN Structures	55
3.6	Conclusion	56
4	Training Algorithms for SICoNNets	58
4.1	Introduction	58
4.2	Overview of Backpropagation Algorithm	61
4.3	Backpropagation Algorithm for SICoNNets	63
4.3.1	Computation of the Sensitivity	63
4.3.2	Computation of the Gradient Terms	67
4.4	A First Order Hybrid Training Method	71
4.4.1	Derivation of the Weight Update from Rprop	72
4.4.2	Derivation of the Adaptive Momentum Rate from Quickprop	73
4.4.3	Derivation of the Adaptive Learning Rate from SuperSAB	74
4.4.4	Thresholds of the Adaptive Parameters	75
4.5	Conjugate Gradient Methods	81
4.5.1	Conjugate Directions	81
4.5.2	Calculation of $\beta(k)$	84
4.5.3	Powell-Beale Restarts	86
4.5.4	Hybrid Conjugate Gradient Methods	86
4.5.5	An Inexact Line Search Method	87
4.5.6	Scaled Conjugate Gradient	90
4.6	Quasi-Newton Methods	92

4.6.1	BFGS Method	93
4.6.2	Modified Versions of BFGS Method	95
4.6.3	One-Step Memory-less BFGS Method	96
4.6.4	Backtracking Line Search Method	96
4.6.5	Combination of a Modified BFGS Update with a Double Dog- leg Method	97
4.7	Levenberg-Marquardt Method	101
4.7.1	Optimized LM Method with Adaptive Momentum	105
4.8	Least Squares Method	106
4.8.1	Integration of LS Method with the Hybrid Training Method (QRPROPLS)	106
4.8.2	Integration of LS Method with the LM Method (LMLS) . . .	108
4.9	Conclusion	109
5	Experimental Analysis of the Training Algorithms	110
5.1	Introduction	110
5.2	Classification Problem	111
5.2.1	Face Database	112
5.2.2	Nonface Database	113
5.3	Image Normalization Techniques	114
5.3.1	Mean Normalization	114
5.3.2	Range Normalization	114
5.3.3	Illumination Gradient Correction	115
5.3.4	Histogram Equalization	115
5.4	Network Structure	116
5.4.1	Activation Functions	118
5.4.2	Comparison Between the Hybrid Training Method, Rprop and Quickprop	121
5.5	Convergence Speed of the Developed Training Algorithms	122
5.6	Classification Accuracy and Generalization Ability	129
5.6.1	K-fold Cross-Validation	129
5.6.2	Classification Results	130
5.7	Conclusion	133
6	Application of SICoNNets	134
6.1	Introduction	134

6.2	A Neural-Based Face Detector	135
6.2.1	Description of the Network Structure	136
6.2.2	Network Training	138
6.2.3	Face Localization Procedure	145
6.2.4	Preliminary Analysis of the Face Detector	147
6.2.5	A Fast Face Detector	149
6.2.6	Evaluation of the Face Detector	152
6.2.7	Orientation Analysis of the Face Detector	156
6.3	A Handwritten Digit Recognition System	159
6.3.1	MNIST Database	159
6.3.2	Network Structure and Training Algorithm	160
6.3.3	Experimental Results and Performance Analysis	160
6.4	Conclusion	163
7	SICoNNets with Rotation Invariance	165
7.1	Introduction	165
7.1.1	Integral Transform Invariants	165
7.1.2	Invariant-Feature Space	167
7.1.3	Invariance by Training	167
7.1.4	Invariance by Structuring the Network and Using Weight Shar- ing	168
7.1.5	Network Specifications for SICoNNet and MLP	170
7.1.6	Rotated Face Database	170
7.1.7	Training and Test Procedures	171
7.2	Experimental Results and Discussion	172
7.3	Conclusion	176
8	Conclusion and Further Work	177
8.1	Introduction	177
8.2	Chapter Summary	178
8.3	Further Directions	180
8.3.1	Extension of the Network Architecture	180
8.3.2	Development of Stochastic Training Algorithms	181
8.3.3	Initialization Procedure	181
8.3.4	Techniques to Improve the Performance of the Applications . .	181

A Pseudo-code of the Training Algorithms	183
B Results of the Face Detection System	190

List of Figures

2.1	A simplified schematic diagram of a biological neuron with its four basic components: dendrites acting as the branching input structure, soma which is the cell body, axon is the connection from the cell body to other neurons through which the output signal is carried away and the synapses is the chemical contacts from neuron to neuron (Source: Microsoft Encarta, 1996).	9
2.2	An artificial model of a simplified biological neuron which simulates its four primitive functions. The biological synapses are represented as synaptic weights and the soma as an adder. The time varying threshold of activity is replaced with an activation function.	10
2.3	A two-layer MLP network structure where the circles represent the artificial neurons, and the lines connecting these neurons are the synaptic weights.	12
2.4	A schematic diagram of the neocognitron architecture with four stages, where each stage consists of a C-layer, S-Layer and V-layer. In each layer the processing cells are arranged into 2-D planes.	16
2.5	The center of the receptive field of each S-cell in the C-plane is located at the corresponding location of the S-cell in the S-plane.	18
2.6	Connections converging to a feature-extracting S-cell. The weights between the C-cells and the S-cells are trainable excitatory weights, whereas the weights between the V-cells to the S-cells are inhibitory weights.	19
2.7	The network structure of LeNet-5 which consists of seven layers.	22
2.8	Location of receptive fields of neighboring neurons in LeNet-5. Each neuron in the feature map has its receptive field overlaps each other by five columns and four rows, and the receptive field does not go beyond the input plane.	23

3.1	A schematic diagram illustrates the application of local receptive fields in the input image or the previous feature map. The receptive field behaves as a filter to extract the same feature across the input image as the all the neuron in the feature share the same set of weight to connect to their receptive fields.	39
3.2	The movement of a receptive field in the input image. Each adjacent shunting inhibitory neuron has its receptive field two units apart. . .	40
3.3	An example of a three layer SICoNNet with an input size of 16×16 pixels. The first hidden layer of the network has two feature maps of size 8×8 , whereas the second hidden layer has four feature maps of size 4×4	41
3.4	Every neuron in the feature map produces an output that is located at the corresponding location of the neuron. Then a sub-sampling operation is applied on the output plane by removing those outputs that are located at odd row or column coordinates of the output plane. Consequently, the size of the output plane is one quarter of the size of the feature feature map.	42
3.5	A local averaging operation is applied on all feature maps of the last hidden layer where four outputs are merged into a single signal that is fed directly to the output neurons.	43
3.6	In a full-connection scheme, each feature map is connected to all feature maps of the succeeding layer.	44
3.7	Partial-connection schemes: (a) a binary connection scheme where each feature map branches out two feature maps in the succeeding layer, and (b) a toeplitz-connection scheme where the feature map has one or more connections with the feature maps in the preceding layer. For example, each feature map of the Layer 2 connects to five feature maps of the Layer 3 and the connections appear along the diagonal of a Toeplitz matrix.	45
3.8	Image of increasing intensities indicating Mach bands at the boundaries between rectangles of different gray levels.	47
3.9	Decision boundaries of a two layer SIANN. Using non-linear activation function for the sigmoid output neuron and different bias parameter b , the network generates different shape of decision boundaries. . . .	52
3.10	A diagrammatic representation of the shunting neuron model. . . .	53

3.11	Input-output transfer characteristics of the enhanced shunting neuron using a logarithmic sigmoid function as activation function. By changing the values of the weights and biases, different shapes of input-output transfer characteristics are generated from the neuron.	54
4.1	A schematic diagram illustrates how the sensitivities from the $L + 1$ layer are back-propagated to form the sensitivity of the i th neuron at the L layer.	62
4.2	When a local averaging operation is applied on the feature map of the last hidden layer, the computed sensitivities of that particular feature map are arranged into a matrix by duplicating each sensitivity into four.	65
4.3	Examples of segmented face patterns taken from images of the AR database.	76
4.4	The curves in the figures represents the number of training epochs required to meet the target MSE while varying the limits of the step size for the CoNN using different size of training sets. In (a), the graphs illustrate the changes in the number of training epochs required to achieve the target MSE when varying the upper limit of the step size of the hybrid training method, and (b) shows the changes in the number of training epochs while varying the lower limit of the step size.	78
4.5	The variations in the number of training epochs as a function of the changes in the limits of the momentum rate of the hybrid method based on five sizes of training sets. In (a), the curves shows the changes in the number of training epochs to reach the stopping criterion while changing the upper limit of the momentum rate and (b) illustrates the convergence of the hybrid training method with respect to the changes in the lower limit of the momentum rate.	79
4.6	Variations in the number of training epochs to reach the target MSE as a function of the upper limit of the adaptive learning rate of the hybrid training method.	80
4.7	An illustration of the line search algorithm developed by Charalambous to compute the optimal step length.	88
4.8	The double dogleg curve: $\vec{W}(k) \rightarrow \vec{W}_{CP}(k) \rightarrow \vec{W}_{\hat{N}} \rightarrow \vec{W}_{NT}(k)$	99

4.9	The search direction of LM algorithm shifts between gradient descent and Gauss-Newton direction by changing the μ parameter. As the parameter μ gets closer to zero, the LM algorithm becomes the Gauss-Newton method, and for large value of μ , the LM algorithm behaves as the gradient descent method with small step sizes.	104
5.1	Examples of face patterns taken from the skin and face detection database created by Phung <i>et al.</i> used for training the developed CoNNs as a face classifier.	112
5.2	Some examples of nonface patterns obtained from the bootstrap training procedure. These nonface patterns are cropped windows from scenery images that have been misclassified by the trained network as faces.	113
5.3	A three-layer binary connected network based on NET-A configuration. At each hidden layer, the number of connections and trainable weights are listed.	118
5.4	The graphs illustrate the respective shape of the four non-linear activation functions. For example, the tansig (hyperbolic tangent) function bounds the output of the neuron between -1 and 1 . Similarly, the logsig (logarithmic sigmoid) function has an output range between 0 and 1	120
5.5	A comparison between the Rprop, Quickprop and hybrid training method in terms of convergence speed. The solid red line represents the convergence curve of the hybrid training method, the dashed blue line is the Rprop method and the dashdot green line is the Quickprop method.	122
5.6	Convergence speed between: (a) LM algorithms with/without a LS method and (b) first order hybrid techniques with/without a LS method. These two figures show that combining a LS method with the training algorithm to further tune the weights of the output layer improves the convergence speed of the training method.	128

6.1	The ROC curves based on networks with different output activation function. Each curve represents the correct detection rate of the network with respect to its false detection rate. The solid black line represents the network with a logarithmic sigmoid output activation function, and the dashed blue and dashdot red lines represent the networks with linear and hyperbolic tangent output activation functions, respectively.	137
6.2	The correct detection rate of the network with respect to its false detection rate when using different network retina sizes ranging from 16×16 to 32×32 pixels, trained and tested: (a) on patterns that are range normalized and (b) on patterns that are histogram equalized and then range normalized between -1 and 1	138
6.3	An example of the image used in the bootstrap training procedure to collect nonface patterns. The image contains a human face where the facial features are removed or distorted. (a) is the original image with a human face, (b) the image is processed to distort the facial features and (c) the facial region of the person is removed from the image.	141
6.4	(a) The changes in the correct classification rates and optimal threshold as a function of the bootstrap session based on: (a) Sung's bootstrap training technique and (b) Garcia's bootstrap training technique. In both techniques, the classification rates for face and non-face increase while more nonface patterns are gradually added to the training set.	142
6.5	The variations in the correct classification rates and the optimal threshold with respect to the bootstrap session based on our modified bootstrap procedure.	144
6.6	Examples of detected face images using: (a) the NNE strategy and (b) the DVS strategy.	148
6.7	(a) A diagram depicts the sub-sampling operation performed on a convolutional map at the final hidden layer of a three-layer network. (b) is the sub-sampled image after rejected all odd rows and columns. (c)-(e) are sub-sampled images taking odd rows and columns.	151
6.8	An output image processed with: (a) sliding step of four pixels and (b) sliding step of two pixels.	152

6.9	Sample of test images from the CMU database where most of the false dismissals are made by the face detector. These test images were scanned from newspapers or photos at low resolution.	155
6.10	Samples of rotated faces used to estimate the rotation angle at which the face detector is accurate. These images have yaw and pitch rotations within the range $\pm 90^\circ$, and $\pm 30^\circ$ for roll rotation.	157
6.11	Classification rate of the developed face detector as a function of the angle of rotation: (a) yaw rotation, (b) pitch rotation, and (c) in-plane rotation.	158
6.12	Samples of the handwritten digit images from the MNIST database. .	160
6.13	Examples of digit patterns in the test set that were misclassified (a) digit four predicted as nine, and (b) digit nine predicted as four. . . .	163
7.1	A training face pattern of size 64×64 taken from the face database created by Phung <i>et al.</i> , rotated in the range $\pm 90^\circ$ at a step of 15° . .	171
7.2	Rotated face patterns from the new face database. These rotated face patterns are obtained by rotating the entire image at different angles and then cropping the faces out of the image.	171
7.3	ROC curves representing the correct detection rates the three developed CoNNs with different receptive field sizes with respect to their false detection rates, tested on face patterns rotated in the complete range of 360°	175

List of Tables

3.1	The connections between the feature maps of Layer 2 and Layer 3 in the toeplitz network architecture. Connections between two hidden layers have the form of a Toeplitz matrix where the connections appear along the diagonal of the matrix.	46
4.1	The upper and lower limits for the momentum rate of the hybrid training method based on the three networks with binary-, toeplitz and full-connection schemes.	80
5.1	Number of trainable weights in a three-layer network architecture with respect to the network configuration. In NET-A configuration, all the neurons in a feature map share the same biases and passive decay rate constant, whereas in NET-B configuration, all the neurons have its own bias and passive decay rate parameters.	117
5.2	Number of connections for the three networks with respect to the developed connection scheme, using receptive field size of 5×5 throughout the network and an input size of 24×24	118
5.3	Non-linear activation functions that are commonly used in artificial neural networks.	119
5.4	Classification rate of a three-layer fully-connected network based on NET-B configuration with various combinations of non-linear activation functions (Ac. Fcns) used in the hidden layers. Different combinations of activation functions in the hidden layers cause the network to yield classification accuracies ranging from 77.5% to 86.5%, with the best performance obtained from combination 5.	121
5.5	Summary of convergence speeds and computation loads of the training algorithms trained on the binary-connected network, based on NET-A and NET-B configurations.	124

5.6	Summary of convergence speeds and computation loads of the training algorithms trained on the toeplitz-connected network, based on NET-A and NET-B configurations.	125
5.7	Summary of convergence speeds and computation loads of the training algorithms trained on the fully-connected network, based on NET-A and NET-B configurations.	126
5.8	Summary of the classification performances of the three developed networks trained by the implemented training algorithms based on NET-A configuration. The performances of the networks are analyzed in terms of correct classification rates for face and nonface and average error rates using a ten-fold cross-validation.	131
5.9	Classification rates and error rates of the three networks trained by the developed training algorithms based on NET-B configuration. . .	132
6.1	Detection performance of the face detector tested on 200 images using the double verification strategy or the normal network strategy. . . .	147
6.2	Number of positive detections and processing times when the exhaustive scanning method or the convolution strategy is used to generate the pyramid of output images from an input image of size 568×686 pixels.	150
6.3	Detection performance of the face detection system based on four large sets of images collected from the Web.	153
6.4	Detection performance of the face detector based on the existing benchmark face databases.	154
6.5	Comparison of the developed face detector with other face detection methods based on the CMU and BioID face databases. The term NP means not reported, CD and FD are the abbreviations for correction detection and false detection, respectively.	155
6.6	Classification rate of the binary-connected network presented as a confusion matrix based on the test set of the MNIST database.	161
6.7	Classification rate of the toeplitz-connected network presented as a confusion matrix based on the test set of the MNIST database.	162
6.8	Classification rate of the fully-connected network presented as a confusion matrix based on the test set of the MNIST database.	162
6.9	Classification performance of the handwritten digit recognition system.	163

7.1	Classification rates of the three CoNNs tested on the rotated face patterns in the range $\pm 90^\circ$ based on different sizes of receptive fields.	173
7.2	Classification rates of the three CoNNs tested on the quasi-frontal face patterns based on different sizes of receptive fields.	173
7.3	Classification performances of the ten best MLPs tested on the rotated face patterns in the range $\pm 90^\circ$. The first five networks have one hidden layer and the next five networks have two hidden layers with different number of neurons.	174
7.4	Classification rates of MLPs with one or two hidden layers, tested on the quasi-frontal face patterns.	175

Acronyms

2-D	2-Dimensional
ANN	Artificial Neural Network
BFGS	Broyen Fletcher Goldfarb and Shanno
CAD	Computer-Aided Diagnosis
CDF	Cumulative Density Function
C-S	Convolutional and Sub-sampling
CoNN	Convolutional Neural Network
DY	Dai and Yuan
DVS	Double Verification Strategy
FCR	Face Classification Rate
FT	Fourier Transform
GLDS	Gray Level Difference Statistic
HS	Hestenes-Stiefel
LM	Levenberg Marquardt
LMS	Least Mean Squares
LS	Least Squares
MACE	Minimum Average Correlation Energy
MSE	Mean Square Error

MLP	Multilayer Perceptrons
MNIST	Modified National Institute of Standard and Technology
NFCR	Nonface Classification Rate
NIST	National Institute of Standard and Technology
NNE	Normal Network Evaluation
OCR	Optical Character Recognition
PDF	Probability Density Function
PR	Polak-Ribière
RBF	Radial Basis Function
ROC	Receiver Characteristic Curve
ROI	Region Of Interest
Rprop	Resilient Backpropagation
SCG	Scale Conjugate Gradient
SGLD	Spatial Gray Level Dependence
SIANN	Shunting Inhibitory Artificial Neural Network
SICNN	Shunting Inhibitory Cellular Neural Network
SICoNNet	Shunting Inhibitory Convolutional Neural Network
SOM	Self Organization Map
SSE	Sum Square Error
SVM	Support Vector Machine
TAP	Target Aim Point
TDNN	Time-Delay Neural Network

Nomenclature

Throughout this thesis, the following mathematical nomenclature has been used to denote the components of the new convolutional neural network architecture and the derivation of its training algorithms.

$\alpha(k)$	step length or the learning rate at the k th iteration
$\Delta\vec{W}(k)$	weight update at the k th iteration
$\delta_{N-1,i}$	sensitivity of the i th neuron in the $(N - 1)$ th layer
ℓ_L	size of the receptive field of the shunting neuron at the L th layer
A	Hessian matrix
G (k)	an approximation to the Hessian at the k th iteration
I	identity matrix
W	a matrix
$ c $	absolute value of the scalar of c
$\ \vec{c}\ $	Euclidean norm or least-square norm of the vector \vec{c}
Ψ_L, Φ_L	activation functions at the L th layer
$\vec{d}(k)$	search direction at the k th iteration
$\vec{g}(k)$	gradient vector of the error function at the k th iteration. The gradient vector is an n -dimensional column vector given by:

$$\vec{g}(k) = [g_1(k), g_2(k), \dots, g_n(k)]^T,$$

where $g_i(k) = \partial E(k)/\partial w_i$ ($i = 1, \dots, n$) is the local gradient

$\vec{V}^T \vec{W}$	the inner product of two vectors
\vec{W}	a vector
$\vec{W}(k)$	a n -dimensional column vector containing all n free parameters (i.e., adaptable weights) of the network at the k th iteration:
$\vec{W}(k) = [w_1(k), w_2(k), \dots, w_n(k)]^T$	
A^T	the transpose of matrix A
A^{-1}	the inverse of matrix A
$a_{L,r}$	passive decay rate constant of the neuron in the r th feature map of the L th layer
$b_{L,r}, d_{L,r}$	bias parameters of the neuron in the r th feature map of the L th layer
$C_{L,k}(x, y)$	excitatory weight at location (x, y) in the receptive field of the shunting neuron in the k th feature map of the L th layer
$D_{L,k}(x, y)$	shunting inhibitory weight at location (x, y) in the receptive field of the shunting neuron in the k th feature map of the L th layer
E, f	cost function or error function
$f'(x)$	the first derivative of function $f(x)$
F_L	size of the feature map at the L th layer
h	activation function at the output layer
n	number of weights in the network
$net_{L,i}$	net input or weighted sum of inputs for the i th neuron in the L th layer
P	number of training patterns
S_I	number of pixels in an image
S_N	number of output neurons
S_T	number of training iterations or epochs

S_{L+1}	number of neurons in the $(L + 1)$ th layer
$\text{sgn}(x)$	the sign of a scalar x
$t_{L,i}^j$	target value of the i th output neuron in the L th layer due to the j th input pattern
$w_{L,ij}$	connection weight from the j th neuron in the $(L - 1)$ th layer to the i th neuron in the L th layer
$Z_{L,i}^j$	output response of the i th output neuron in the L th layer due to the j th input pattern
$\ast\ast\ast$	2D convolution operator