



شبكة المعلومات الجامعية
التوثيق الإلكتروني والميكروفيلم

بسم الله الرحمن الرحيم



MONA MAGHRABY



شبكة المعلومات الجامعية
التوثيق الإلكتروني والميكروفيلم



شبكة المعلومات الجامعية التوثيق الإلكتروني والميكروفيلم



MONA MAGHRABY



شبكة المعلومات الجامعية
التوثيق الإلكتروني والميكروفيلم

جامعة عين شمس

التوثيق الإلكتروني والميكروفيلم

قسم

نقسم بالله العظيم أن المادة التي تم توثيقها وتسجيلها
علي هذه الأقراص المدمجة قد أعدت دون أية تغيرات



يجب أن

تحفظ هذه الأقراص المدمجة بعيدا عن الغبار



MONA MAGHRABY



AIN SHAMS UNIVERSITY
Faculty of Computer
& Information Sciences
Computer Science Department

Improving Convolutional Neural Networks Learning Through Adaptation

Thesis submitted to the Department of Computer Science

Faculty of Computer and Information Sciences

Ain Shams University, Egypt

**In partial fulfillment of the requirements for the master's degree of
Computer and Information Sciences**

Submitted by:

Zainab Mohamed Fouad Ibrahim

B.Sc. of Computer Science,

**Teaching Assistant, Department of Computer Science,
Faculty of Computer and Information Sciences, Ain Shams University.**

Under Supervision of:

Prof. Dr. Abdel-Badeeh Mohamed Salem

Professor of Computer Science

Faculty of Computer and Information Sciences, Ain Shams University.

Prof. Dr. Mohamed Ismail Roushdy

Professor of Computer Science

Faculty of Computer and Information Sciences, Ain Shams University.

Dean of Faculty of Computer and Information Technology, Future University in Egypt.

Dr. Marco Alfonse Tawfik

Lecturer in the Department of Computer Science,

Faculty of Computer and Information Sciences, Ain Shams University.

Table Of Contents

	Page
Acknowledgment	iii
Abstract	iv
List of Publications	vi
List of Figures	vii
List of Tables	viii
List of Algorithms	ix
List of Abbreviations	x
1 Introduction	
1.1 Thesis Motivation.....	3
1.2 Thesis Objectives.....	3
1.3 Thesis Methodology.....	3
1.4 Thesis Contribution.....	4
1.5 Thesis Organization.....	4
2 Overview of Convolution Neural Networks	
2.1 Pattern recognition background	6
2.2 Machine learning background.....	10
2.2.1 ML and deep learning.....	10
2.2.2 ML and AI.....	11
2.2.3 How ML works.....	11
2.2.4 ML methods.....	12
2.3 CNN Structure.....	13
2.4 CNN Architectures.....	21
2.5 CNN Applications.....	22
3 Related Works	
3.1 Automatic design of CNN model.....	25
3.2 CNN enhancement methods	27
3.3 CNN hyper-parameters optimization methods.....	40
3.4 Discussion.....	45

4	Hyper-Parameter Optimization of Convolutional Neural Network Based on Particle Swarm Optimization Algorithm	
4.1	Swarm Intelligent Algorithms	48
4.2	CNN integration with particle swarm optimization algorithm method.....	51
5	Experiments and Results	
5.1	Experimental Methodology.....	54
5.2	Experimental Results	56
6	Conclusions and Future Work	
6.1	Conclusions	59
6.2	Future Work	60
	References	61

Acknowledgment

Thanks merciful GOD for all your gifts in my whole life. I would like to thank my supervisors Professor Dr. Abdel-badeeh Salem, Professor Dr. Mohamed Roushdy, and Dr. Marco Alfonse who really supported and guided me greatly through the development of my Master's and offered me their precious time with experience and advice.

Thanks to all my colleagues and all my friends at the Department of Computer Science.

Finally, with all my appreciation and love that no words can express, I dedicate this Master to my dear parents, my beloved husband Dr. Islam, and my sons Ammar and Youssef who always provided me with love, prayers, blessings, advice and care. They are behind any success in my life.

Abstract

Recently, Convolution Neural Network (CNN) has accomplished great success in numerous issues of machine learning. Many machine learning methods have been developed for such objectives, for example, Artificial Neural Network (ANN), logistic regression, Support Vector Machine (SVM), deep learning, etc.

Deep learning (specifically CNN) is one of the strategies by which can delude the challenges of the feature extraction process. Usually, deep learning models are capable of extracting the proper features by themselves. Also, deep CNN models are usually designed manually and the key parameters of it are decided by experience and repeated tests which incredibly limit the applications of deep CNN.

Therefore, it is a great challenge to design the proper deep CNN model and reduce the dependence on manual involvement and expertise. So, this thesis will discuss the improvement of the convolution neural network design from different aspects with various methods. The CNN improvements include how to automatically design CNN model without operator intervention, change on convolution or pooling layers, adding some features to save the computational resources, and how to use adaptive and optimized CNN parameters.

On the other side, many hyper-parameters of the CNN can affect the model performance. These parameters are depth of the network, numbers of convolutional layers, numbers of kernels with their sizes. Therefore, it may be a huge challenge to design an appropriate CNN model that uses optimized hyper-parameters.

In this thesis, a design architecture method for CNNs is proposed by utilization of Particle Swarm Optimization (PSO) algorithm to learn the optimal CNN hyper-parameters values.

In the experiment, MNIST (Modified National Institute of Standards and Technology) database of handwritten digit recognition was used. The experiments appear that the proposed approach can find an architecture that is competitive to the state-of-the-art models with a testing error of 0.87%.

List of Publications

1. Zainab Fouad, Marco Alfonse, Mohamed Roushdy and Abdel-Badeeh M. Salem, “Improving the Design of Convolution Neural Network Architecture: Enhancement Methods, Applications and Challenges”, proceeding of the 10th International Conference on Applied Internet and Information Technologies AIIT 2020, pp. 289-295. (Indexed in Scopus)
2. Zainab Fouad, Marco Alfonse, Mohamed Roushdy and Abdel-Badeeh M. Salem, “Hyper-parameter optimization of convolutional neural network based on particle swarm optimization algorithm”, Journal of Bulletin of Electrical Engineering and Informatics, Vol 10, No 6: December 2021, pp. 3377-3384. (Indexed in Scopus)
3. Zainab Fouad, Marco Alfonse, Mohamed Roushdy and Abdel-Badeeh M. Salem, “Survey of Convolutional neural network hyper-parameters optimization methods”, International Journal of Intelligent Computing and Information Sciences. (Accepted)

List of Figures

Figure	Page
Figure 2.1 English pattern.....	6
Figure 2.2 The root of plants and human stomach patterns.....	7
Figure 2.3 DNA and Protein pattern.....	7
Figure 2.4 Speech signals.....	8
Figure 2.5 EGK signal.....	8
Figure 2.6 Texture patterns.....	9
Figure 2.7 Faces and Fingerprints patterns.....	9
Figure 2.8 CNN Example	14
Figure 2.9 The building blocks of CNNs.....	15
Figure 2.10 Convolution Step.....	16
Figure 2.11 The more convolution layers will get more complex features.....	17
Figure 2.12 Depth in convolution step.....	17
Figure 2.13 The ReLU operation	18
Figure 2.14 Max Pooling.....	19
Figure 2.15 Fully Connected Layer	20
Figure 3.1 Automatic design of CNN model	26
Figure 3.2 The model of incremental learning	30
Figure 3.3 The NIN model.....	31
Figure 3.4 The distinction between straight convolutional layer and mlpconv layer	31
Figure 3.5 The dynamic filter module	32
Figure 3.6 Example of the downside of max and average pooling.....	33
Figure 3.7 The ADCNN model	34
Figure 3.8 The Multi-Scale Convolutional Networks Architecture.....	35
Figure 3.9 The Spatial transformer module	36
Figure 3.10 The flowchart of Albelwi and Mahmood framework.....	42
Figure 4.1 Proposed method flowchart.....	51
Figure 5.1 Sample of MINST database	54
Figure 5.2 Training accuracy during the optimization process.....	56

List of Tables

Table		Page
Table 2.1	Recent CNN applications.....	22
Table 3.1	Automatic design of CNN Model	27
Table 3.2	CNN Enhancement Methods.....	37
Table 3.3	CNN hyperparameter optimization methods.....	44
Table 5.1	Experiments parameters	55
Table 5.2	Comparison of MINST classification error of state of art methods.	57

List of Algorithms

Algorithm	Page
Algorithm 4.1 Artificial Bee Colony	49
Algorithm 4.2 Particle Swarm Optimization.....	50
Algorithm 4.3 CNN integration with particle swarm optimization.....	52

List of Abbreviations

Abbreviation	Definition
ABC	Artificial Bee Colony
ACNN	Adaptive Convolutional Neural Network
ACO	Ant Colony Optimization
ADCNN	Adaptive Deep Convolutional Neural Network
AFSO	Artificial Fish Swarm Optimization
ANN	Artificial Neural Network
BA	Bat Algorithm
BFO	Bacterial Foraging Optimization
CGP	Cartesian Genetic Programming
CNN	Convolutional Neural Network
DPSO	Distributed PSO
FA	Firefly Algorithm
GA	Genetic Algorithms
HBMO	Honeybee Mating Optimization
HDNN	Hybrid Deep CNN
HENet	Highly Efficient Convolutional Neural Networks
HORD	Hyper-Parameter Optimization via RBF and Dynamic coordinate search
LDWPSO	linearly decreasing weight PSO
MAE	Mean Absolute Error

MINST	Modified National Institute of Standards and Technology
ML	Machine Learning
MSE	Mean Square Error
NIN	Network in Network
NMM	Nelder-Mead Method
PSO	Particle Swarm Optimization
RBF	Radial Basis Function
ReLU	Rectified Linear Unit
SaCNN	Scale adaptive CNN
SI	Swarm Intelligence
SVM	Support Vector Machine