



AIN SHAMS UNIVERSITY
FACULTY OF ENGINEERING
Computer Engineering and Systems

Automatic Recognition of Arabic Handwriting using Hybrid Intelligent Networks

A Thesis submitted in partial fulfillment of the requirements of
Doctor of Philosophy in Electrical Engineering
(Computer and Systems Engineering)
by

Taraggy Mohiy Ghanim
Masters of Science
(Computer Engineering and Systems Department)
Faculty of Engineering, Ain Shams University, 2012

Supervised By
Prof. Hazem M. Abbas
Prof. Mahmoud I. Khalil

Date: 2020



AIN SHAMS UNIVERSITY
FACULTY OF ENGINEERING
Computer Engineering and Systems

Automatic Recognition of Arabic Handwriting
using
Hybrid Intelligent
Networks

Examiners'
Committee

Name and affiliation

Signature

Prof.Nawal Ahmed El-Fishawy
Computer Engineering and Systems Department
Faculty of Engineering, Menoufia University.

.....

Prof.Hoda Korashy
Computer Engineering and Systems Department
Faculty of Engineering, Ain Shams University.

.....

Prof.Hazem M. Abbas
Computer Engineering and Systems Department
Faculty of Engineering, Ain Shams University.

.....

Prof.Mahmoud I Khalil
Computer Engineering and Systems Department
Faculty of Engineering, Ain Shams University.

.....

Date:2020

Statement

This thesis is submitted as a partial fulfillment of Doctor of Philosophy in Electrical Engineering, Faculty of Engineering, Ain Shams University. The author carried out the work included in this thesis, and no part of it has been submitted for a degree or a qualification at any other scientific entity.

Taraggy Mohiy Ghanim

Signature

.....

Date: 2020

Researcher Data

Name: Taraggy Mohiy Ghanim

Date of Birth: 19/6/1984

Place of Birth: Cairo, Egypt

Last academic degree: Master of Science

Field of specialization: Pattern Recognition, Artificial Intelligence

University issued the degree : Ain Shams University

Date of issued degree : 2012

Current job : Assistant Lecturer in Misr International University

Thesis Summary

Key words: Arabic Handwriting Recognition, Random Forest, Kullback-Leibler Divergence, Pyramid Histogram of Gradient, Support Vector Machine, Agglomerative Hierarchical clustering, Deep Convolutional Neural Network

Automatic Recognition of Arabic Handwriting is a pervasive field that has many challenging complications to solve. Such complications include big databases and complex computing activities. Chapter 1 introduces our motivation and challenges, while chapter 2 presents our related work. Our approach is a multi-stage cascading system, and is proposed in Chapter 3. It is based on applying hybrid machine learning techniques consecutively. Hybrid cascading recognition systems aim to improve the learning ability and increase recognition rates. The approach stages start with data-mining which is essentially needed to work effectively on big databases. Agglomerative hierarchical clustering technique is followed to split the database into partially inter-related clusters for the data mining process. Each test image is matched to one cluster. Cluster members are then ranked in an ascending order based on our new proposed ranking algorithm. This ranking algorithm begins with computing Pyramid Histogram of oriented Gradients (PHoG), followed by measuring divergence by the Kullback-Leibler method. Finally, the classification process is applied only on the highly ranked matching classes, to assign a class membership to each test image. Adjusting the classification process to only consider the highly ranked database classes supports classification and enhances the overall performance.

Recently, the field of computer vision and building recognition systems is being switched to the concept of deep learning to serve the field of feature engineering. Deep learning techniques demonstrated efficiency in building a better performing machine learning model. This thesis assesses the effect of six different deep Convolution Neural Networks (DCNNs) architectures in the field of offline Arabic handwriting recognition. Different neural networks' architectures and design aspects are applicable in a very challenging manner.

Chapter 4 displays our experiments. The approach is tested on the IFN-ENIT Arabic database. Enhanced classification results are achieved relative to recent proposed systems. Chapter 5 discusses the complexity analysis.

Acknowledgment

Contents

I would like to express my sincere gratitude to my supervisors for their continuous support and guidance. Besides my supervisors, I would like to thank the rest of my thesis committee. Also, I would like to thank my family: my parents, my husband and my children for supporting me spiritually throughout my life.

Taraggy Mohiy Ghanim
Computer Engineering and Systems
Faculty of Engineering
Ain Shams University
Cairo, Egypt
May 2020

Acknowledgment

Contents

Contents	xi
List of Figures	xiv
List of Tables	xvi
1 Introduction	1
1.1 Aim of the Research	1
1.2 Motivation and Challenges	2
1.2.1 Computer Vision and Automatic Handwriting Recognition	2
1.2.2 Online Recognition versus Offline Recognition	2
1.2.3 Arabic Language Challenges	2
1.2.4 Effect of Hybrid Recognition Systems on Arabic Handwriting Recognition	3
1.2.5 Effect of Deep learning in the field of Feature Engineering	4
1.3 Publications Resulting from Research	4
1.4 Original Contributions of Thesis	5
1.5 Outline of the Thesis	6
2 Literature Review	7
2.1 Concepts and Analysis	7
2.2 Categorization based on Segmentation Approaches	8
2.3 Features Extraction	8
2.4 Integration of classifiers	9
2.5 Post-Processing Support	10
2.6 Hidden Markovian Models (HMM) based approaches	10
2.7 Support Vector Machines (SVM) based approaches	12
2.8 Neural Networks (NN) based approaches	13
2.8.1 Summary of Chapter, Conclusions and Results	16
3 Proposed Approach	18
3.1 Stage 1: Nearest Neighbor Set Construction stage	20
3.1.1 Preprocessing and Segmentation	21
3.1.2 Regional and Geometric Features Extraction	22
3.1.3 Random Forest for Feature Selection	25
3.1.4 Clustering	26
3.1.4.1 Agglomerative Hierarchical Clustering	26
3.2 Stage 2: Ranking Stage (PHoG and KL-divergence)	27

List of Figures

3.2.1	Statistical Features Extraction: Pyramid Histogram of Gradients (PHoG)	28
3.2.2	Divergence Measure: Kullback-Leibler (KL)	32
3.3	Stage 3: Classification Stage	34
3.3.1	Classification using Multi-class Support Vector Machines	35
3.3.2	Classification using Convolutional Neural Network	36
3.3.2.1	CNN Different Layers	36
3.3.2.2	CNN different Architectures	38
3.3.2.2.1	LeNet-5:	39
3.3.2.2.2	AlexNet:	40
3.3.2.2.3	VGG-16:	40
3.3.2.2.4	Inception (GoogLeNet)	41
3.3.2.2.5	ResNet	43
3.3.2.2.6	ResNeXt	44
3.3.2.2.7	DenseNet	44
3.4	Summary	46
4	Results and Discussions	47
4.1	Database	47
4.2	Matching Process Sensitivity	48
4.3	The ranking stage	52
4.4	The classification stage	55
4.5	Comparison with similar systems	56
4.6	Experiments on individual sets using SVM	61
4.7	Convolution Neural Networks (CNN)	62
4.8	Reasons of misclassification	66
5	Complexity Analysis	67
5.1	Complexity Analysis of Agglomerative Hierarchical Clustering	68
5.2	Complexity Analysis of Pyramid of Histogram of Oriented Gradients PHoG	68
5.3	Complexity Analysis of Support Vector Machines (SVM)	69
5.4	Complexity Analysis of Convolutional Neural Network (CNN)	69
6	Conclusions and Future Work	71
6.1	Conclusions	71
6.2	Future Work	73

List of Figures

1.1 Different Categories of Arabic Characters:(A) Arabic Characters do not join to their left neighbor, (B) stand-alone character, (C) Different dots number and positions	3
3.1 The proposed system overview.	19
3.2 Stage 1: (n is number of test images, m is total number of database classes, L is total number of classes per cluster	21
3.3 A sample image during preprocessing	22
3.4 Universe of Discourse.	22
3.5 A sample image after computing number of holes and PAWs.	23
3.6 Eccentricity, F1 and F2 are the two foci of the ellipse	23
3.7 Orientation θ of word bounding ellipse.	24
3.8 Concavity features.	25
3.9 Feature Importance	25
3.10 The Clustering Process	27
3.11 Stage 2: L is the cluster size, W is number of high rank classes	28
3.12 Pyramid Histogram of Gradients (F: frequency, G: oriented gradient)	29
3.13 Effect of Number of Orientation Bins on True Positive Rate (TPR).	31
3.14 Convolution Example with Stride=2	37
3.15 Max Pooling with a 2x2 Filter and Stride = 2	38
3.16 CNN Architectures	39
3.17 LeNet-5 Architectures [105]	39
3.18 AlexNet Architectures [98]	40
3.19 VGG-16 CNN architecture [16]	41
3.20 Inception Network Primary Design [17]	42
3.21 Inception Cell	42
3.22 Replacing Big Spatial Filters by Successive Smaller Ones	43
3.23 Residual Block	43
3.24 ResNet34 versus ResNet50 versus ResNeXt	45
3.25 DenseNet Block	46
4.1 Nearest Neighbor Set Selection Sensitivity	49
4.2 Different set sizes and their frequency.	49
4.3 Effect of different features on Sensitivity	50
4.4 Effect of number of training samples on sensitivity.	51
4.5 The Effect of Ranking on the Expected Success Rate	53
4.6 Sample of Clusters and Early Ranks of Correct matching Class	54
4.7 PHoG and Geometric features with SVM.	55

List of Tables

4.8	The error rates of using SVM versus KNN.	57
4.9	Character Modeling in HMM	58
4.10	Effect of Nearest Neighbors Set Length on SVM Recognition Rate.	61
4.11	AlexNet: Achieved accuracy and testing time versus learning rate	62
4.12	VGG-16:: Achieved accuracy and testing time versus learning rate	63
4.13	GoogleNet: Achieved accuracy and testing time versus learning rate	64
4.14	Res50 Net: Achieved accuracy and testing time versus learning rate	64
4.15	ResNeXt Net: Achieved accuracy and testing time versus learning rate	65
4.16	DenseNet: Achieved accuracy and testing time versus learning rate	65
4.17	Samples of Misclassified Images	66
5.1	Complexity Evaluation on IFN-ENIT Dataset	70