Al-Quds University

Faculty of Graduate Studies

Arabic Alphabet Deaf Sign Gestures Recognition Based on Deep Machine Learning Methods

Ohood Adel Salameh Darabee

M.Sc. Thesis

Jerusalem-Palestine

1440-2019

Arabic Alphabet Deaf Sign Gestures Recognition Based on Deep Machine Learning Methods

Prepared by:

Ohood Adel Salameh Darabee

B.Sc. Computer Science

Hebron University/Palestine

Supervisor: Dr. Abdullah Kamal

A thesis submitted in fulfillment of requirements for the degree of Master of Computer Science of Al-Quds University.

1440/2019

Al-Quds University

Deanship of Graduate Studies

Computer Science

Thesis Approval

Arabic Alphabet Deaf Sign Gestures Recognition Based on Deep Machine Learning Methods

Student Name: Ohood Adel Salameh Darabee

Registration No: 21311351

Supervisor: Dr. Abdullah Kamal

Master thesis submitted and accepted date: 13 / 7 /2019

The names and signatures of the examining committee members are as follows:

1-Head of committee: Dr. Abdullah Kamal signature:**.....**

2- Internal Examiner: Dr. Badie Sartawi signature: **.....**

3- External Examinar: Dr.Ahmad Hasasneh signature: **.....**

Jerusalem- Palestine

1440/2019

Dedication

I dedicate this to souls in the sky; may god rest them in Heaven.

I dedicate this to my mother and family.

I dedicate this to my supportive husband and my only army; Lutfi

Thanks for all

Declaration

I certify that this thesis submitted for the degree of Master of Science is the result of my own research, except where otherwise acknowledged, and that this thesis (or any part of the same) has not been submitted for a higher degree to any other university or institution.

Signed ……..

Ohood Adel Salameh Darabee

Date: 13 / 7 / 2019

Acknowledgments

I would like to give my sincere thanks and gratitude to people that give assistant and support for me during master research.

Special mention for Dr Abdullah Kamal and all my doctors who inspired me to work in the field of recognition and classification. Thank you for your invaluable comments and suggestions.

I would like to give special thanks for the support and love of my husband, who supported me financially and morally.

I would like to thank Palestinian Red Crescent Society. Special thanks for the sign specialist Ms. Engy Alabed.

I would like to give special thanks to the Media institute at Al-Quds University. Special thanks for the photographer Mr.Shadi Barakat

Finally yet importantly thanks goes to Al-Quds University, especially for Faculty of high studies\ Computer science department.

Abstract

Sign language continue to be the best method to communicate between the deaf and hearing impaired. Hand gestures enable communication between deaf people during their daily lives rather than speaking. In our society, Arabic Sign Language (ArSL) is only known for deaf people and specialist, which makes the community of deaf people narrow. Recognizing and documentation of ASL have only been paid attention recently. This thesis proposes a real time gesture recognition system for Arabic sign language based on Restricted Boltzman Machines (RBMs) following by the use of tiny images. RBMs have the ability to code images as structures of limited features taken from a large alphabet. This process continue to repeat itself in a deep networking to get an efficient sparse representation of the initial data in the feature space. A complex problem of recognition in the input space is thus transformed into an easier one in the feature space. In this thesis, we show that ASL can thus be achieved using tiny images instead of conventional Bag-of-Words (BoWs) methods. After appropriate coding, a simple classifier in the feature space suffices to compute the probability of a particular character according to the input image.

Table of Contents

Declaration I

Acknowledgments II

Abstract III

Contents IV

List of Figures VI

List of Tables VIII

List of Acronyms IX

1. Chapter 1 1

1.1. Introduction 1

1.2. Problem definition and overview 2

1.3. Introduction to deaf sign language 3

1.4. Thesis objectives 6

1.5. Major contributions 7

1.6. Thesis benefits 7

1.7. Thesis constraints 7

1.8. Thesis organization 8

2. Chapter 2 9

Background and related work 9

2.1. Introduction to Gesture Recognition 9

2.2. **Introduction to ArSL** 10

2.3. Current approaches to ArSL 11

2.3.1. Classification methods 12

2.4 Deep learning approaches 29

2.4.1 Restricted Boltzmann Machines 30

2.4.2. RBM learning algorithm 33

2.4.3. Deep Belief Network 34

2.5 Summary and recommendations 35

3. Chapter 3 36

Proposed Model 36

3.1. Introduction 36

3.2. The Database 38

3.3. Image Pre-processing 39

3.3.1. Image Grey Scaling.....39
3.3.2. Use of Tiny Images 40
3.3.3. Data Whitening and Normalization 41
3.4. Unsupervised Feature Space Construction......42
3.5. Summary.....51
4. Chapter 4.....52
Results.....52
4.1. Introduction.....52
4.2. Data Collecting Results.....53
4.3. Image pre-processing results.....54
4.4 Features extraction based on RBM training on tiny images.....58
4.6Summary 65
5.Chapter 5.....66
Conclusion and Future Work. 66
5.1: **Introduction**.....66
5.2: Contribution.....66
5.3: Results:.....66
5.4: Limitation......66
5.5: Future Work.....66
References 68

List of Figures

Figure 1—1 : Sign language recognition stages 2
Figure 1 2—Using iconic features from nature to represent words.[8] 4
Figure 1 3—Using miming in relating to space to describe prepositions.[8] 4
Figure 1—4 : Arabic sign language alphabets 5
Figure 0—1 Arabic sign language alphabets 11
Figure 0 : 2—K-Nearest Neighbor Algorithm example.[38] 15
Figure 0—3 Probabilistic parameters of a hidden Markov model (an example), where X represents the states, y represents possible observations, a represents the state transition probabilities, and b represents the output probabilities 18
Figure 04— HMM Chain to predict the weather (Bishop, 2006). 19
Figure 05— : HMM Chain for weather prediction (Bishop, 2006). 19
Figure 0—6 : High-level structure of an SVM approach. 20
Figure 0—7 : linear SVM. 21
Figure 0—8 : non-linear SVM. 21
Figure 0—9 : SVM mapping. 22
Figure 0—10 : Soft max linear classification for two different classes 23
Figure 0—11 : Logistic sigmoid activation function: with 24
Figure 0—12 : An example for scalar product of neural networks. 27
Figure 0—13 : Basic Structure of multilayer neural network. 27
Figure 0—14 : Left figure represent a classical Boltzmann machine, where there is a fully connection between the visible units themselves, hidden units themselves, and between the visible and hidden units. Right figure represent an RBM where there is no direct conn 31
Figure 0—15 : Restricted Boltzmann Machines (RBM layer wise) structure that build Deep Belief Network. 34
Figure 31—: the proposed model 37
Figure 32—: Arabic Alphabet sign language Database used in the proposed model 39
Figure 33—: The difference between colored and grey scaled image 40
Figure 34—: Resizing image from the normal size to 100*100 pixel then to 42*42 pixel. Notice that the tiny image still recognizable 41

Figure 35—: On the left is normalized 8*8-patch sample, on the right whitened 8*8-sample patch.....42
Figure 36—: shows the training of RBM machine using constructive divergence learning 44
Figure 01—: The Arabic Alphabet signs Dataset done by the specialist Ms. Engy Alaabed 54
Figure 02—: Samples from the training dataset (أ،ب،ت) taken under different attribution.....54
Figure 03—: Samples of image cropping to eliminate the un-necessary elements in the image and centered the hand gesture. 55
Figure 04—: Samples of image grey scaled 56
Figure 05—: Samples of image resizing on the dataset.....56
Figure 06— Applying data normalization on the Arabic Alphabet database 57
Figure 0—7: the 1024 features extracted by RBM training on the dataset. 59

List of Tables

Table 4.1	Classification Result 1	page 59
Table 4.2	Classification Result 2	page 60
Table 4.3	Classification Result 3	page 60
Table 4.4	Comparison of ArSlR systems Result	page 62

List of Acronyms

ASL	**Arabic Sign Language**
RBMs	**Restricted Boltzmann Machines**
BoW	**Bag of Words**
ArSLR	**Arabic Sign Language Recognition**
ROI	**Region of interests**
DBNs	**Beep Belief Networks**
CD	**Constructive Divergence**
SVM	**Support Vector Machine**
HCI	**Human Computer Interface**
K-NN	**K-Nearest Neighbor**
HMM	**Hidden Marcov Model**
NN	**Neural Network**
MLP	**Multi-layer perceptron**
ICA	**Independent Component Analysis**
DBN	**Deep belief Network**
RBM	**Restricted Boltzmann Machine**
MCMC	**Marcov Chain Monte-Carlo**

1. Chapter 1

Introduction

1.1. Introduction

The most natural way that human being used to communicate with each other is by using voice, gestures, and human-machine interfaces. However, the last method is still very primitive and force to adapt the machine requirements. Instead of using voice, hand gestures enable communication between deaf people during their daily live. However, Arabic Sign Language (ArSL) or gesture recognition is known only for deaf people and specialists in our society, thus the community of deaf people is very limited and narrow. Recently, sign language recognition systems for American, British, Indian, Chinese, Turkish, and many international sign languages have received much attention compared to the Arabic sign language. Therefore, developing an Arabic sign language recognition system (ArSLR) is needed. A review of recent developments in sign language recognition can be found in [1][2][3].

As shown in figure 1, ArSLR can be performed through two main phases: detection and classification. In the detection phase, each captured image is pre-processed, improved, and then the Regions of Interest (ROI) is identified using a segmentation algorithm. The output of the segmentation process can thus be used to perform the classification process. Indeed, accuracy and speed of detection play an important role in obtaining accurate and fast recognition process. In the recognition phase, a set of features are extracted from each segmented hand sign and then used to perform the recognition process. These features can therefore be used as a reference to understand the differences among the different classes. As mentioned earlier, ArSLR systems have only been paid attention recently, where few attempts have investigated and addressed this problem, see for example [4]–[6]. Therefore, the question of ArSL recognition is still an open question and need to be investigated using new and different machine learning methods. This thesis proposes a new Arabic sign

recognition system based on new machine learning methods and a direct use of tiny images.

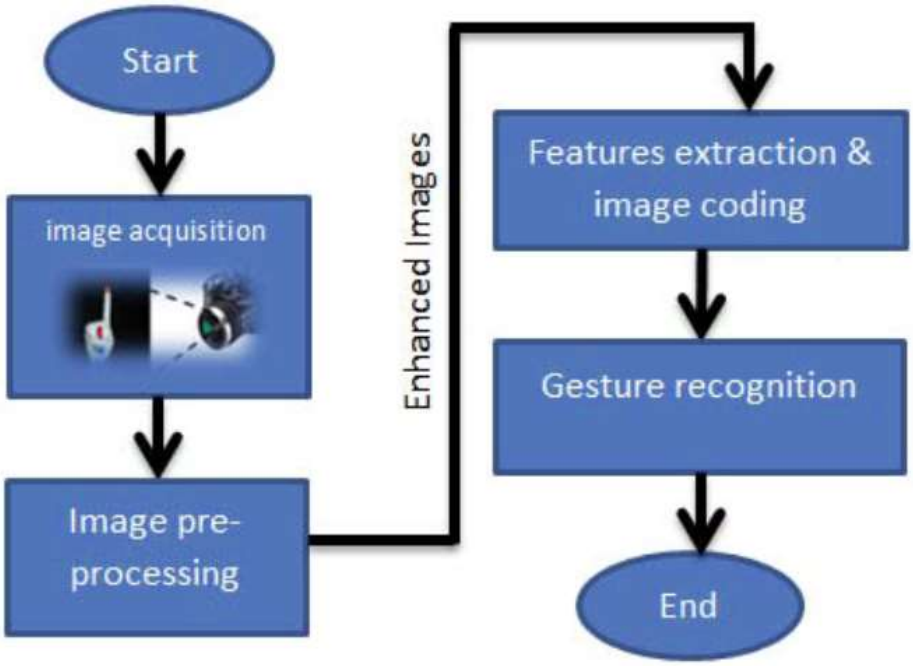

Figure 1—1 : Sign language recognition stages

In this introduction, we start by identifying the research problem statement and the substantial motivations to propose a new approach to ArSLR problem. We then describe the main objectives, advantages, and contributions to develop this work. We end this chapter by presenting an outline of the remaining chapters of this thesis.

1.2. Problem definition and overview

Deaf people use sign language to communicate with each other, but they cannot communicate with normal people because they cannot read sign language, which makes the community of deaf people very narrow. Statistics show that over 3% of the palestinian populations are hearing impaired [1]. In particular, according to the Palestinian Central Bureau of Statistics, 19% of the disabled Palestinian people are deaf and mute [7]. Indeed, helping those people is very important and thus developing systems capable of translating sign languages into text or spoken language is highly needed. Developing such systems will definitely participate in facilitating the communication between the hearing impaired and the normal people. Many approaches have been proposed to achieve sign language recognition; however, few