Republic of Iraq
Ministry of Higher Education
and Scientific Research
University of Babylon
College of Information Technology
Software Department

HAND GESTURE RECOGNITION BASED ON MULTI-CONNECT ARCHITECTURE ASSOCIATIVE MEMORY

A Thesis
Submitted to the Council of the College of Information Technology
University of Babylon in Partial Fulfilment of The Requirements For The Degree
of Doctor of Philosophy In Information Technology-Software

By
NOOR FADEL HUSSEIN SULIMAN

Supervised By
PROF. DR. EMAD ISSA ABDUL KAREEM

2023 A.D. 1445 A.H.

بِسْمِ اللَّهِ الرَّحْمَٰنِ الرَّحِيمِ

﴿ نَرْفَعُ دَرَجَاتٍ مَّن نَّشَاءُ ۗ وَفَوْقَ كُلِّ ذِي عِلْمٍ عَلِيمٌ ﴾

صدق الله العظيم

سورة يوسف الاية ... 76

Supervisor Certification

I certify that this thesis **Hand Gesture Recognition Based on Multi-Connect Architecture Associative Memory** has been prepared under my supervision at the department of Software / College of Information Technology/ University of Babylon, by **Noor Fadel Husseain** as a partial fulfillment of the requirements of the degree of Doctor of Philosophy in Information Technology-Software

Signature:

Name: Prof. Dr. Emad Issa Abdul Kareem

Date: 15 / 10 / 2023

The Head of the Department Certification

In view of the available recommendations, I forward this thesis for debate by the examination committee.

Signature:

Name: Prof. Dr. Ahmad Saleem Abbas

Date: / /2023

Dedication

To the Majesty on the throne of motherhood "my mother",

To my father dear God save you,

To the person who supports me owner to wonderful heart, my husband,

My children's Narges and Ali

My brothers,

My sisters,

all one in my life

With all my gratitude

ACKNOWLEDGEMENTS

In the Name of Allah, the Most Merciful, the Most Compassionate all praise be to Allah, and prayers and peace be upon Mohamed His servant and messenger

All the praises and thanks are to **Allah** for his graces that enabled me to accomplish the requirements of my study.
My deepest thanks and gratitude to **my parents** for their love and support throughout my life. thank you both for giving me strength,
My dear husband, my family, thank you from the depth of my heart.

My sincere thanks to my supervisor **Dr. Emad Issa Abdul Kareem** for his guidance, support, motivation and encouragement throughout the period of this work which was carried out

Finally I would like to thank all the kind, helpful and lovely people who helped me directly or indirectly to complete this work and I apologize for not being able to mention all the names.

Noor (2023)

Declaration

The work in this thesis, **Hand Gesture Recognition Based on Multi-Connect Architecture Associative Memory**, is original and no portion of the work referred to here has been submitted in support of an application for another degree or qualification of this or any other university or institution of learning.

Signature:

Date: 10 / 10 /2023

Noor Fadel Hussein

Index

List of Contents	**II**
List of Figures	**IV**
List of Tables	**VI**
List of Algorithms	**VII**
List of Abbreviations	**VII**
List of Publication	**IX**

List of Contents

Chapter One		
1.1	Introduction	1
1.2	Related Works	2
1.2.1	Review of Hand Gesture Recognition Based on machine learning Techniques	2
1.2.2	Review of Hand Gesture Recognition Based on ANN And DNN Techniques	6
1.3	Problem Statement	10
1.4	Main Goal and Objectives	10
1.5	Scope of Research	11
1.6	Contribution	11
1.7	Challenges	12
1.8	Organization of Thesis	13
Chapter Two		
2.1	Introduction	14
2.1.1	Hand Gesture Methods	14
2.2	Applications of Gesture Recognition System	16
2.3	Hand Gesture Recognition	17
2.3.1	Hand Gesture Recognition Based on Machine Learning Techniques	18
2.3.2	Computer Vision Techniques for Hand Gesture Recognition	20
2.3.3	Sign Language Recognition	22
2.4	Hand Detection and Segmentation	24
2.4.1	Hand detection-based machine learning	24
2.4.2	Hand Detection based Deep learning	27
2.5	Associative Memory	30
2.6	Sign Language Recognition System Evaluation	38
2.7	Summary	40

Chapter Three		
3.1	Introduction	41
3.2	Research Methodology	41
3.3	The Proposed System	43
3.3.1	Hand Detect and Segmentation	46
3.3.2	Recognition of Hand Gestures	50
3.3.3	Translate Sign Language Into Text	59
3.4	Summary	60
Chapter Four		
4.1	Introduction	61
4.2	Case Study	63
4.3	Real Time Hand Gesture Recognition Evaluation	64
4.3.1	Accuracy for Hand Gestures Recognition	65
4. 3.2	RTHGR Efficiency	66
4.3.3	Evaluation of The Proposed System in Real-Time.	68
4.4	Qualitative Evaluation	71
4.4.1	The Curriculum of The Study	71
4.4.2	Study Community	71
4.4.3	The Study Sample	72
4.4.4	Study tool	72
4.4.5	Questionnaire Components	73
4.5.6	Staging Scale	75
4.5	Results of The Field Study	76
4.5.1	Statistical Description of The Study Sample	76
4.5.2	Study Sample According to The Gender Variable	76
4.5.3	The Study Sample According To The Knowledge of Sign Language	77
4.5.4	A Study Sample According To The Variable Job Title	78
4.5.5	Study Sample According To Years of Service Variable	78
4.5.6	Statistical Description of the RTHGR System	79

4.5.6.1	Second Pivot: Interactive Interface of The System.	79
4.5.6.2	The Third Pivot: The Proposed System for Learning About Sign Language	81
4.5.6.3	The Fourth Pivot: The Accuracy of The System.	83
4.5.6.4	The Fifth Pivot: Decision-Making Speed.	85
4.5.6.5	The Sixth Pivot: Cost of The Hardware	86
4.5.6.6	The Seventh Pivot: of Applicability	88
4.6	Summary	91
Chapter Five		
5.1	**Conclusion**	92
5.2	**Future Works**	94

List of Figures

Chapter One		
1.1	Hand Gesture Detection and Foreground Segmentation Stages	3
1.2	Hand gesture recognition based on hand topology	5
Chapter Two		
2.1	Hand Gestures. (a) Glove-based attached sensor (b) computer vision–based camera	16
2.2	Application of Hand Gesture	17
2.3	Simple Example on Deep Learning Convolutional Neural Network Architecture	20
2.4	Using Computer Vision Techniques to Identify Gestures	21
2.5	Computer Vision for Hand Gesture Recognition	21
2.6	RGB vs. HSV Colour Space	24
2.7	Rustle of Hand Gesture Detection (Segmentation)Techniques	26
2.8	21 Different Landmarks on The Hand	28
2.9	Landmark Point for Hand Detection	29

2.10	The Landmarks of a Hand	30
2.11	A General Block Diagram of An Associative Memory	31
2.12	The Multi-Connect Architecture (MCA) Associative Memory	33
2.13	All the Possible Vectors with Their Weight Matrices and Majority Description	34
2.14	Confusion Matrix for Multiclass	39
Chapter Three		
3.1	Proposed System Methodology Flow	42
3.2	Real-Time Hand Gestures Recognition Model (RTHGR)	45
3.3	The Result of Hand Detection Operation	49
3.4	The RTHGR Learning Phase.	51
3.5	A Lookup Table for The Training Hand Gesture	53
3.6	The RTHGR Converge Phases	55
3.7	Compute The 8 Possible Vectors, smd And weight For The Pattern.	58
Chapter Four		
4.1	Case Study Steps	63
4.2	A Letter "A and B" From American Sign Language	64
4.3	Accuracy For Hand Gestures Recognition	66
4.4	Recognize the signs with convergence rate	68
4.5	population according to the job title.	71
4.6	Distribution of Sample by Place of Work.	72
4.7	Study Sample According to The Gender Variable	76
4.8	Study Sample According to Knowledge of Sign Language	77
4.9	Sample According to Years of Service Variable.	78
4.10	Rate of Answers on The Second Pivot	79
4.11	Rate of Answers on The Third Pivot	81

4.12	Rate of Answers on The Fourth Pivot	83
4.13	Rate of Answers on The Fifth Pivot	85
4.14	Rate of Answers on The Sixth Pivot	87
4.15	Rate of Answers on The Seven Pivot	89

List of Tables

1.1	Summary of Review of Hand Gesture Recognition Based on AI Techniques.	6
1.2	Summary of Review of Hand Gesture Recognition Based on deep learning.	9
2.1	The energy function values corresponding to each vector	38
4.1	The Result of The Proposed Method in Real-Time.	69
4.2	Comparison of The Proposed Method With Other Methods	70
4.3	Relative Weight Scale	76
4.4	Analysis of Pivot of Interactive Interface.	80
4.5	Analysis of Pivot of the proposed system for learning about sign language	81
4.6	Analysis of Pivot of the accuracy of the system	83
4.7	Analysis of Pivot of decision-making speed	85
4.8	Analysis cost of pivot of the hardware requirements	87
4.9	Analysis of Pivot of applicability	89

List of Algorithms

Algorithm 3.1: Hand Detection and Segmentation	**48**
Algorithm 3.2: Hand Gestures Trainer	**52**
Algorithm 3.3: Hand Gestures Detector	**57**
Algorithm 3.4: Hand Gestures translator	**60**

List of Abbreviations

Acronym	Meaning
2DCNN	2-D Convolutional Neural Networks
3DCNN	Three Dimension Convolutional Neural Networks
ANN	Artificial Neural Networks
AR	Augmented Reality
Arsl	Arabic Sign Language
ASL	American Sign Language
BSL	Brazilian Sign Language
ChSL	Chinese Sign Language
CNN	Convolutional Neural Networks
ConvLSTM	Convolutional Long-Short-Term-Memory
Cp	Converge Pattern
CV	Converge Vector
DC-CNN	Double-Channel Convolutional Neural Network
DNN	Deep Neural Networks
EMG	Electromyography
FCM	Fuzzy C Mean
HCI	Human-Computer Interaction
HGR	Hand Gesture Recognition
HOG	Histogram of Gradient
HSV	Hue, Saturation, And Value
ISL	Indian Sign Language
KNN	K -Nearest Neighbour
KSL	Korean Sign Language
LMs	Landmark Points
LSTM	Long- Short Term Memory

MCA	Multi-Connect Associative Memory
Md	Majority Description
ML	Machine Learning
NUI	Natural User Interface
PCA	Principal Component Analysis
RC	Remote Control
RGB	Red, Green, And Blue
RNN	Recurrent Neural Network
ROI	Region Of Interest
RTHGR	Real-Time Hand Gesture Recognition
Smd	Sum Majority Description
SVM	Support Vector Machine
Svw	Stored Vector's Weight
tmd	tested majority description.
tvw	Tested Vector's Weight
VR	Virtual Reality

List of Publications

The first

Paper : Detecting Hand Gestures Using Machine Learning Techniques

The Journal: Ingénierie des Systèmes d'Information (ISI)

CiteScore 2022: 2.7

Publisher: Lavoisier, France

The link : https://www.iieta.org/journals/isi/paper/10.18280/isi.270612

The second

Paper: Computer Vision Techniques for Hand Gesture Recognition: Survey

The Journal : Communications in Computer and Information Science

publisher : **Springer Nature,** Germany

The link : https://link.springer.com/chapter/10.1007/978-3-031-35442-7_4

The third

Paper: Real -Time Hand Gesture Recognition Based on Multi-Connect Architecture Associative Memory in Human Computer Interaction

Accepted in 4th International Conference on Pure Sciences (ICPS)

The Journal : AIP Conference Proceedings

Publisher: American Institute of Physics, United States

ABSTRACT

Recognizing hand gestures under computer vision is crucial for facilitating communication between the deaf and dumb community and normal people, as well as between the elderly who cannot wear hand gloves or sensors. interpretation of hand gestures is a fundamental part of Human-Computer Interaction that relies heavily on the ability to understand and respond to human signs.

Recognition of highly complex, similarly shaped hand gestures is critical. As a result, there is an urgent need for a system with accuracy and the ability to recognize these gestures, as the problem of inaccurately recognizing hand gestures may hurt the human community that relies on gestures to transmit their desires. Given the importance of applications for hand gesture recognition, a completely new method has been developed for recognizing hand gestures and applying it to the hand gesture data set.

The proposed system goal is to illuminate the most crucial steps in the hand gesture detection process. which is the process of detection and recognition of hand gestures by using hand landmark points for detecting hands and using the Multi-Connect Architecture (MCA) of associative memory as a new trend in the recognition phase. The problem with the similarity between the signs is because of the strong correlation between the movements of the fingers. In addition to the non-high accuracy between complex and very similar signs (e.g., A, S, and E). And the problem of response time in hand gesture recognition in real-time, Thus, using the MCA associative memory neural network improved efficiency in dealing with the correlation between similar patterns by taking similar vectors for each hand gesture pattern only once. The proposed system gave high accuracy in real-time via using the confusion matrix and showed promising outcomes with accuracy for

American, Chinese, and Arabic sign language and numbers at 95.42%, 92.13%, 93.55%, and 94%, respectively. As well as the system's work in an uncontrolled environment, in addition to being an applicable system by converting the sign into its meaning as words and sentences, not just letters.

CHAPTER ONE

GENERAL INTRODUCTION

1.1 Introduction

Hand gesture recognition (HGR) is an important and major research topic in the field of assistive technology. The hand gesture is a convenient way to transfer information as well as an alternative to devices such as the mouse and keyboard. Also, it helps older adults who cannot walk or talk communicate with their caregivers when they need help through hand gestures [1].

According to the notion of human-computer interaction (HCI) technology, humans and machines may interact naturally. Conventional human-machine contact is mostly done through devices like mice, keyboards, and displays. It goes without saying that these devices generally need to be connected to a computer. In some cases, such as virtual reality (VR), remote control (RC), and augmented reality (AR), the procedure of linking these devices is insufficient. As a result, it is critical to do research on how to create an HCI environment that is in tune with human communication behaviors [1,2].

The use of hand gestures varies from one application to another, depending on the user's cultural background, application domain, and environment. Likewise, with expressive gestures, sign language. One of the new paradigms today is the so-called natural user interface (NUI). The basic idea is that the user does not have to rely on additional hardware as a means of input or output and can operate the computer system in the same way that real objects would [3].

Recognizing hand gestures has gained increasing importance in recent times for two main reasons: first, the growth of the deaf and hard-of-hearing population around the world, with a small percentage, and second, the development of applications based on vision and touchless control on devices such as video games, smart TV control, and virtual reality applications [4].

Hand gestures provide a non-physical link to the computer for user comfort and safety, as well as the ability to handle complicated and virtual settings in a much easier manner than traditional ways in a significant variety of HCI applications . Hand gesture applications, on the other hand, need the ability to correctly apply and comprehend various gestures [5].

Previous researches shows that the deaf and dumb community will increase by 2050 compared to what currently exists due to noise and other reasons, and with the increase in the number of the deaf and dumb community compared to the ignorance of ordinary people in sign language and the need of the two communities to communicate, recognition of sign language has become of great importance by using technological methods [2].

1.2 Related Worke

Recognizing hand gestures is an important and vital topic, and previous work by researchers in this section has been addressed in two parts, based on machine learning methods and deep learning methods, within the concept of computer vision only.

1.2.1 Review of Hand Gesture Recognition Based on Machen Learning Techniques

This technique [6] is implemented by reading a frame as an image and then extracting it manually using the YCbCr color space filter. Then it is converted to a black-and-white image. The experiment was conducted using 180 random hand gesture frames taken from random people (see Figure. 1.1).

It was found that under-regulated lighting circumstances, high-resolution cameras, and a short distance away, gestures were examined with three criteria that directly affect the identification rate, namely noise, light levels, and hand size. The method conducts finger counting and gesture recognition based on the maximum distance between the observed fingers, with an average performance of 50–70% of successful detection for 14 movements based on color space-based Y–Cb–Cr segmentation. The Y–Cb–Cr color space was shown to be helpful.

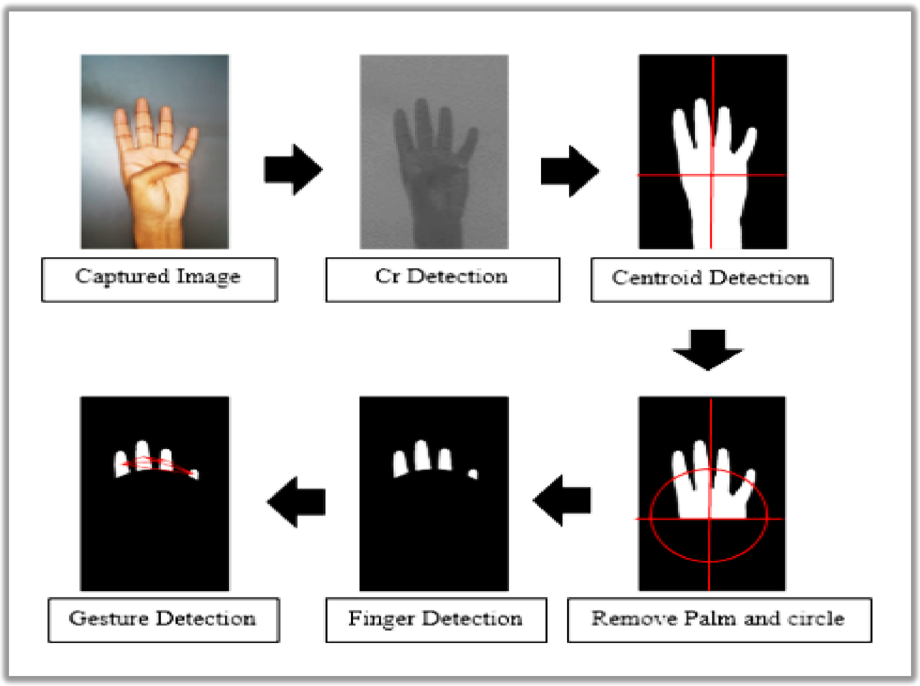


Figure 1.1: Hand Gesture Detection and Foreground Segmentation Stages [6]

The method that was proposed in [7] first processed the images after extracting a histogram of gradient (HOG) features from the input data that was provided by a portable webcam. The portable webcam provides the proposed system with the input data, which comprises four different hand motions to be interpreted. After the image acquired has been processed using the data entered, the graph of the vector gradient features will be retrieved from the obtained image. Following this, the processed image is looked up in a database that contains images of gestures. The KNN method is used to perform the comparison and recognition of the image. The slide show is then controlled by the image that was initially detected.

American Sign Language was used in the study [8] Images are collected from the webcam and preprocessed. The figure is divided into two parts by means of polygon approximations and convex decomposition approximations. Features are extracted by recording the unique features between the different convex parts of the hand. The resulting singularity was used as the extracted feature vector. This training includes features that are almost unique to various hand gestures.

The work that was presented in reference number [9] identified detections by employing a hybrid of the k-nearest neighbor (KNN) and decision tree classification methods for hand gesture recognition. The classification of [10] [11] gestures was also accomplished with the assistance of a Support Vector Machine (SVM).

A three-part gesture recognition robot system was introduced [12] an intelligent robot system, gesture recognition, and an Android app. Three independent discriminators contained in the gesture recognition system were verified to control gesture recognition based on the use of an algorithm based on feature extraction and template matching. The average rate of gesture recognition was 94.6%.

Principal Component Analysis (PCA) and classifiers such as the Support Vector Machine (SVM) and k-near-neighbor (kNN) were used in addition to the recognition experiments. Reducing this similarity between frames is an important step in pattern recognition. The method was applied to the Irish Sign Language data set [13].

This article [14] focuses on a dynamic 3D gesture recognition system that uses hand location information. It uses the natural structure of the hand topology (later called hand skeletal data) to extract effective manual kinematic descriptors from a set of gestures. The descriptors are then encoded into statistical and temporal representations using the Fisher kernel and a multi-level time hierarchy, respectively. The proposed approach was applied to a three-handed gesture dataset, each containing 10, 14, and 25 gestures. Also, the proposed evaluation is from the perspective of hand gesture recognition and low-latency gesture recognition see (Figure 1.2).

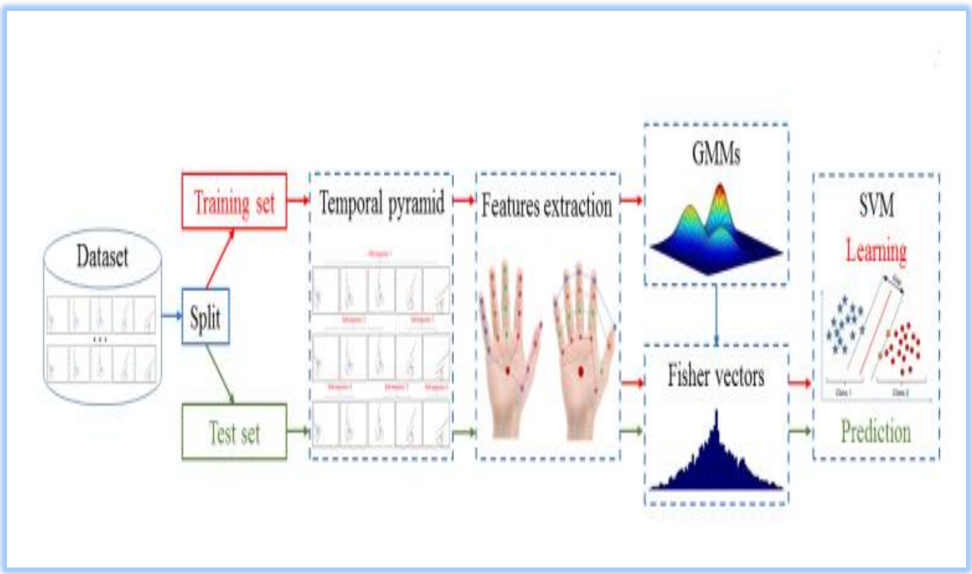


Figure 1.2: Hand gesture recognition based on hand topology [14].

In general, and by reviewing previous research on hand gesture recognition, the general review can be summarized in addition to the details mentioned in (Table 1.1).

Table 1.1: Summary of Review of The Hand Gesture Recognition Based on AI Techniques.

Ref.	Year	Segmentation	Feature extraction	Recogntion	Accuracy	Application Area
[6]	2018	Y–Cb–Cr	maximum distance between fingers	-	50 to 70	HCI
[7]	2018	-	HOG	KNN	effectiveness	HCI , Power point control
[15]	2019	YUV & CAMShift	-	naïve Bayes	high	human - machine interaction
[16]	2018	local neighbor method	convex hull detection	-	96%	human–robot interaction
[17]	2018	bounding box & hand mask	-	-	High	Interaction with deformable object & tracking
[18]	2018	YUV	-	naïve Bayes cl	High	human and machine system

1.2.2 Review of Hand Gesture Recognition Based on ANN And DNN Techniques

Even if a related dataset is used to pre-train the network and take advantage of transfer learning, one of the primary challenges associated with deep learning-based image recognition is the requirement for a vast amount of labeled data. In addition, there are other issues, such as design optimization, which can be difficult and time-consuming [19]. These challenges are one of the main problems with deep learning-based image recognition.

[20] This research presents the design of the Arabic Sign Language Recognition Model. There were four phases of the recognition process involved

in turning the alphabet into letters: There are three stages: an initial image loading phase in which Arabic sign language letter images are loaded for use in subsequent training and testing phases a Pre-processing phase, wherein essential features for completing the recognition are extracted via image processing techniques such as Normalization, Resizing, Image Augmentation, and Filter; CNN and other methods of deep learning are used during the training phase, whereas the testing phase shows how well the model performs in practice.

[21] A three-dimensional convolutional neural network model for learning region-based spatiotemporal aspects of hand movements has been proposed. This model takes as its input a sequence of RGB frames that were taken by a straightforward camera. In order to extract the features from the full video sample, a 3DCNN was used, and a SoftMax layer was utilized in order to perform the classification. It acquired recognition rates of 84.38 percent, 34.9 percent, and 70 percent on the three datasets, respectively, by using three gesture datasets derived from color videos: KSU-SSL, ARSL, and ASL classes.

The [22] study presents a double-channel convolutional neural network (DC-CNN) in which the original image is preprocessed using canny detection to detect the edge of the hand before being given to the network. To categorize output results, each two-channel CNN has its own weight and SoftMax classifier.

The study [23] provided an effective, fast, and easy method for dynamic hand gesture recognition based on two-line features extracted from two-line real-time video. The feature selection was used to represent the hand shape of the Kurdish Sign Language's dynamic word recognition. Features extracted in real-time from preprocessed manual objects were represented by optimized values for the captured binary frames. Finally, used an artificial neural network classifier

MLP to recognize hand gestures. The accuracy of this research is 98 for the recognition of 10 words.

Graph CNNs are employed in the study of [24] to create a 3D network of the hand surface. Also, it used Graph CNN to deal with three dimensions based on Landmark points, and the method obtained an accuracy of 92, Hand segmentation and 3D hand position estimation, including skeletal constraints, can be learned using a CNN-based approach developed by [25]. Deep learning models using parallel convolutional neural networks (CNNs) to process the positions of hand skeleton joints have been introduced in [26] for the purpose of 3D hand gesture detection. The accuracy for the recognition 14 gesture is 91.28 and the accuracy for the recognition 28 gesture is 84.35.

Another suggested method In [27] is to track hands using webcams. used skin color detection and morphology in the detection stage for 6 hand gestures used in home-control and then used a deep convolutional neural network with recognition rates for training and test data of 99.90% and 95.61%, respectively.

A new method based on the Deep CNN feeds resized images directly to the network, ignoring segmentation and recognition levels and directly classifying hand gestures. The system operates in real-time, with a result of 97.1% for simple backgrounds and 85.3% for complex backgrounds[28].

Used Convolutional Neural Network (CNN) in another research paper[29], where the original images are preprocessed to detect the hand before being transmitted to the network and the output is classified using the SoftMax classifier. The detection rate of the suggested system is 90 for 28 Arabic sign languages.

In this paper [30], an improved recurrent neural network is proposed for skeletal-based dynamic hand gesture recognition. After finger movement and general movement features are extracted, they are fed into a bidirectional recurrent neural network (RNN) along with the extracted finger movement features to describe finger movements. This method is effective and superior to other art initiation methods.

This study [31] proposes combining the capabilities of two deep learning technologies, a convolutional neural network (CNN) and a recurrent neural network (RNN), to automatically recognize hand gestures using depth and structure data. The purpose of this study is to use depth data and apply CNNs to extract important spatial information from images. The CNN+RNN can recognize different gestures with an accuracy of 85.46%. It combines both skeletal and depth information to extract temporal and spatial information.

A proposed deep learning-based method for temporary 3D shape recognition issues [32] is primarily based on a combination of a convolutional neural network (CNN) and a long short-term memory (LSTM) recurrent network. Introducing a two-level education method that focuses on CNN education at first and then changes the whole method (CNN + LSTM) at the second level.

Table 1.2. Summary of Review of Hand Gesture Recognition Based on deep learning.

Ref.	Year	Segmentation	Feature extraction	Recogntion	Data set	Accuracy	Application Area
[22]	2020	Gaussian Mixture model	CNN	-	hand gestures	95.96%	hand gesture recognition for human
[26]	2018	hand-skeletal joints'	-	CNN	-	91.28% 84.35%	HCI
[27]	2019	-	features extraction by CNN	ADCNN	hand gestures	99%	HCI people - communicate

[28]	2019	–	deep CNN	-	hand gestures	simple background 97.1% complex background 85.3%"	Control a piece of consumer devices
[33]	2019	skin color detection and morphology	deep CNN	-	hand gestures	95.61%"	Home -control
[34]	2019	skin color	CNN	SVM	hand gestures	98.52%	HCI

1.3 Problem Statement

A comprehensive analysis of the computer vision-based method for hand gesture detection reveals the following issues:

1. The visual analysis of hand motions is complex and highly structured, making hand gesture recognition a complex task. There are a number of gestures that are complex and similar to each other, which leads to gesture misidentification (especially in the case of a closed hand).
2. The problem with the similarity between the signs is because of the strong correlation between the movements of the fingers. In addition to the non-high accuracy between complex and very similar signs (e.g., A, S, and E).
3. The problem of response time in sign language recognition in real-time.

1.4 Main Goal and Objective

The main aim of the proposed system is to establish an RTHGR model for recognizing hand gestures in real-time. The following aims are outlined in order to accomplish this:

1. Using associative memory techniques with the Real-Time Hand Gesture Recognition (RTHGR) model. Improving the reliability of the proposed

system that classified gestures that are complex and similar to each other by using Multi-Connect Architecture (MCA) associative memory techniques.

2. Using good methods to detect hand gestures within various states, like illumination variation, clutter, and scaling, based on the shapes of hands by using landmark points.
3. To evaluate the RTHGR model, more than one data set has been used.
4. Solving the deaf and dumb communication problem by facilitating the conversion of the signs into words that reflect its meaning.

1.5 Scope of Research

The proposed system focuses on one of the recognition technologies, gesture recognition systems. A Real-Time Hand Gesture Recognition (RTHGR) model is developed to determine the different meanings of sign language (i.e., numbers, letters, and others). Applying the proposed model to more than one type of sign language (American, Arabic, and Chinese) as the dataset within the scope of the proposed system.

The proposed system results are very important in the translation process between dumb people and normal people. Thus, the translation process as a whole sentence or set of words is within the scope of the proposed system.

1.6 Contribution

The proposed system a Real-Time Hand Gesture Recognition (RTHGR) that recognizes many gestures, the contributions can be summarized as follows:

1. Using associative memory to detect different hand gestures as a new trend.

2. Develop an efficient hand gesture recognition system with high accuracy via using the MCA associative memory neural network in real-time, and applying it to more than one dataset.
3. Improving the efficiency in dealing with the correlation between similar patterns based on using MCA associative memory neural network by taking similar vectors for each pattern only once.
4. Converting the meaning of the signs as words and sentences. By combining the letters of words resulting from sign recognition to build a sentence

1.7 Challenges

There are several challenges, as follows:

1. The scales of the hand vary in size. Clutter in the background that blends with the skin colour is another obstacle. There are a number of obstacles to consider while dealing with the color of the skin, such as lighting changes.
2. Movement is hands-on, precise, and complex. So capturing the gestures is quite a challenge (the closed hand).
3. Some gestures are difficult to reproduce for the same person.
4. Not a universal language. There is no uniform sign language used around the world that differs from one country to another.

1.8 Organization of Thesis

The thesis has the following outline:

1. The theoretical underpinnings are discussed in Chapter 2. The article provides context for the Proposed system being described.
2. The methods employed in this Proposed system to accomplish its aims are discussed in Chapter 3.
3. The outcomes of the trials are discussed and analyzed in Chapter 4. In addition to discussing the themes and results of the questionnaire.
4. The conclusion and relevant future work for this Proposed system are presented in Chapter 5.

CHAPTER TWO

THEORETICAL FOUNDATION

2.1 INTRODUCTION

This chapter presents a review of hand gesture methods, the application of hand gestures, hand gesture recognition and sign language, and associative memory concepts with their types. In addition, hand detection methods and multi-connect architecture associative memory are also explained in this chapter.

2.1.1 Hand Gesture Methods

The goal of gesture recognition is to make a system that can recognize human gestures to transfer data or give commands. It detects human hand motion and translates it into commands [6]. To collect data or information for the identification of hand gestures. This may be accomplished primarily using two techniques [5]:

a) Sensor-Based Approaches

Collecting data consisting of finger and hand position, movement, and trajectory requires the use of sensors or devices physically attached to the user's arm or hand. The main sensor-based methods are:

1. The glove-based approach measures hand and finger position, acceleration, degrees of freedom, and flexion.

2. Electromyography (EMG) measures the electrical impulses of human muscles and decodes biological signals to detect the movement of fingers.

3. Other uses include mechanical, ultrasonic, electromagnetic, and other tactile techniques.

b) Vision-Based Approaches

With these approaches [8], human movements are captured by one or more cameras [9], and vision-based devices can handle many features to interpret gestures (i.e., color, texture, and shape of the hand). The sensor does not have this characteristic. Although these approaches are simple, they can pose many challenges. Systems that require a variety of lighting, complex backgrounds, the presence of objects of hand-like skin color (clutter), and some criteria such as detection time, speed, durability, and computational efficiency [8] [9]. The vision-based approach differs in many issues:

1. Camera numbers.
2. Customer requirements.
3. Speed and reaction time
4. Low-level highlights are used (histogram, outline, edge, etc.).
5. Structure of the environment (simple or complex environment, development speed, and lighting conditions).

The term "vision-based" is most commonly used to capture images and recordings of bare hands without gloves or markers [10].

1. The sensor-based approach reduces the need for preprocessing and pitch assembly, which are the basis of traditional vision-based motion detection systems.
2. The vision-based approach does not require gadgets (other than cameras), but it requires a huge amount of information to prepare a framework that can be generalized to invisible scenarios [3] and [7].
3. Vision-based approaches eliminate the need for a device (besides a camera), but they need large amounts of data to

train systems that can generalize to unseen scenarios see (Figure 2.1).

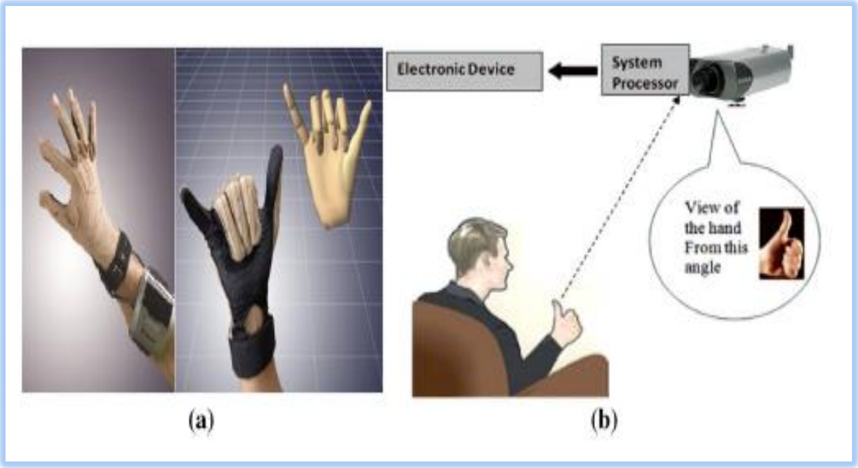


Figure 2.1: Hand Gestures. (**A**) Glove-Based Attached Sensor (**B**) Computer Vision–Based Camera [3].

2.2 Applications of Gesture Recognition System

Gesture recognition is a new interactive technology that makes human-machine interactions more natural, convenient, and efficient. The gesture recognition applications are as follows:

1. Talking to the computer instead of clicking the mouse can add or move images by simply flicking the hand [11].
2. Medical surgery: Gestures can be used to control the visual display and assist users with disabilities as part of rehabilitation therapy [2].
3. Gesture-based game control [13].
4. Hand gestures to control home appliances such as MP3 players and TVs, etc. [4].
5. Sign language applications are very important in order to bridge the communication gap between deaf people and people who do not

suffer from speech or hearing problems [3]. There are many sign languages, most notably American, British, Hindi, Arabic, etc.

6. The driver can be seen in the eye to choose a radio station, and you can change the station without looking away from the road by simply flicking the hand [6].
7. When the term "immersion" is used, it is frequently associated with augmented reality [12]. Immersive media can create an environment that can be stopped and interacted with. "Immersion" is a term that has attracted much attention in technology see (Figure 2.2).

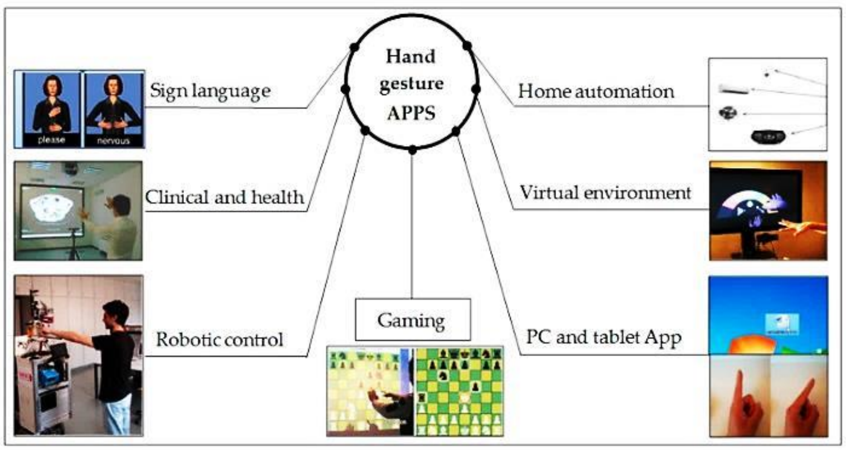


Figure 2.1: Application of Hand Gesture [13].

The next subsections provide a literature survey related to hand gesture systems. This survey highlights different hand gestures recognition that has been developed using different techniques to improve their recognition.

2.3 Hand Gesture Recognition

The topic of hand gesture recognition is vast, and much research has been done in recent years. This section surveys the most recent studies on hand gesture recognition. will also contrast the various methodologies and applications described in the examined papers. The review indicated that most articles use a single camera (webcam or laptop) for the data acquisition process [35].

In recent years, vision-based hand gesture recognition research has shifted to integrating more in-depth information sources like Microsoft Kinect and Leap Motion Controller [3].

2.3.1 Hand Gesture Recognition Based on Machine Learning Techniques

The topic of hand gesture recognition is vast, and much research has been done in recent years.

Artificial intelligence offers a good and reliable technique used in a wide range of modern applications because of using a learning role principle [36]. This section reviews gesture segmentation methods, feature extraction methods used to images to abstract those features, and hand gesture classification algorithms.

1. **Segmentation**

Image segmentation is one of the most frequently techniques used in image processing. The appropriate method for segmenting images, in fact, depends largely on the quality of the images and the dataset, if it is static or dynamic hand motion. Many methods of segmentation have been dealt with in recognizing hand gestures, starting from discovering edges, color spaces, clustering, using threshold or contour methods [37].

2. **Feature Extraction**

The accuracy of the quality of any algorithm depends on the quality of extracted features, because these features, or rather important information, give a complete impression of the classification or recognition process. Therefore, accurate information extraction from images must aid in the process of gesture recognition [8].

The Fourier descriptor, and fingertips, contour hand, motion gesture, or centroid, capture the basic structure. It can be considered a common method of feature extraction [36].

3. Classification

There are many classification algorithms that can be used to identify different gestures, whether the gesture is dynamic, sign language, etc. The extracted features are sent to one of the classification methods, and the data is divided for testing and training the classification algorithm, through our survey of a number of researchers [7].

There is no doubt that understanding hand gestures of all kinds, whether sign language or expressing a specific matter, is a very important topic of research for researchers, especially for the many applications in which hand gestures are involved, as they help deaf and dumb people communicate, as well as the elderly in the health care sector, robotics, and other fields. As a result, the motivation for the research that dealt with the recognition of hand gestures based on vision is an important topic to know the latest technologies used, and the possibility of using these technologies in many vital applications makes communication with computers more natural. The (Figure 2.3) below shows an example of using the principle of deep learning with gesture recognition.

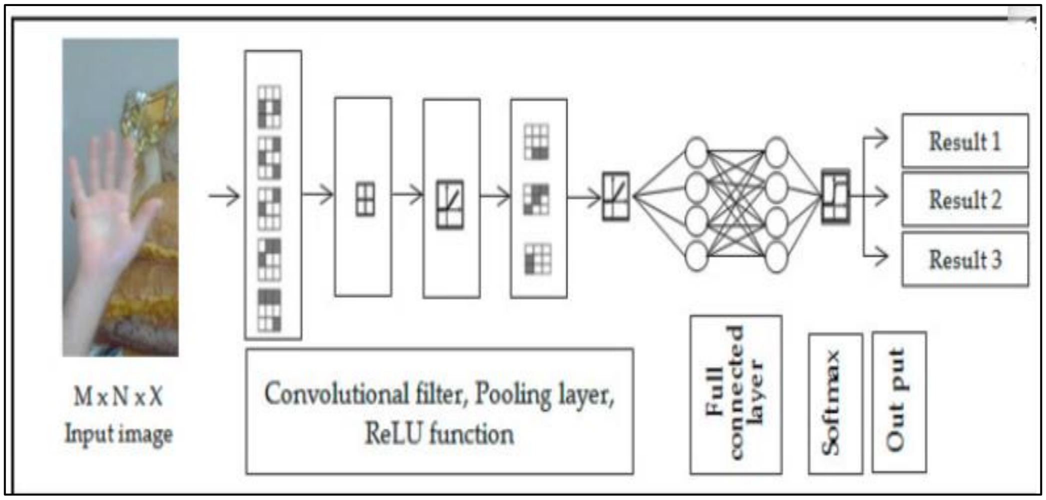


Figure 2.3: Simple Example on Deep Learning Convolutional Neural Network Architecture [27].

2.3.2 Computer Vision Techniques for Hand Gesture Recognition

It is very important to pay attention to new techniques that deal with the idea of recognizing gestures using computer vision. Hand gesture recognition has recently emerged as a critical component of the human-computer interaction (HCI) concept, allowing computers to capture and interpret hand gestures. In addition to their use in many medical applications, communication between the hearing impaired, device automation, and robot control, hand gestures are of particular importance as a form of nonverbal communication.

Because it allows for communication between humans and computers without the requirement of physical contact [16], the camera vision-based is a technique that is widely used, acceptable, and applicable. There are several different types of camera setups that can be used, including, fisheye, monocular, TOF, and IR [20]. However, this method presents a number of difficulties, such as inconsistent lighting, problems with the background, effects such as occlusions, a complicated background, time to process that must be traded off against resolution and frame rate, and foreground or background items that present a comparable skin tone or otherwise appear to be a hand [17,21]. Figure (2.4) provides an easy explanation of the camera vision-based that is utilized for deciphering and understanding various hand movements

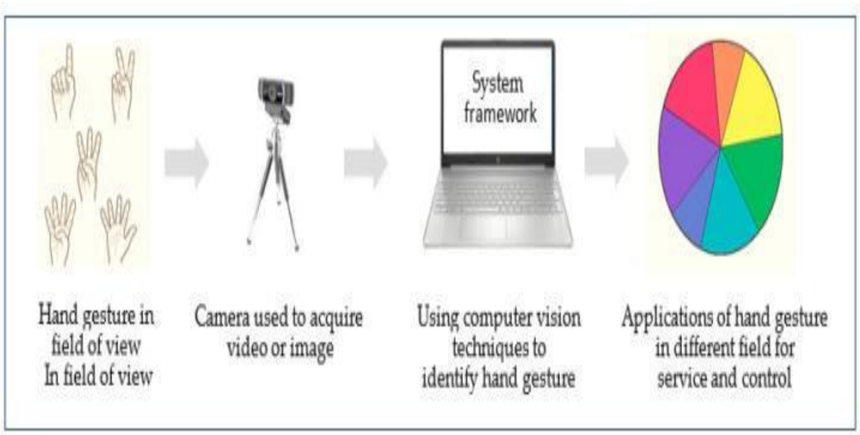


Figure 2.4: Using Computer Vision Techniques to Identify Gestures [2].

Many technologies are based on artificial intelligence and deep learning The seven most commonly used technologies can be summarized in (Figure. 2.5).

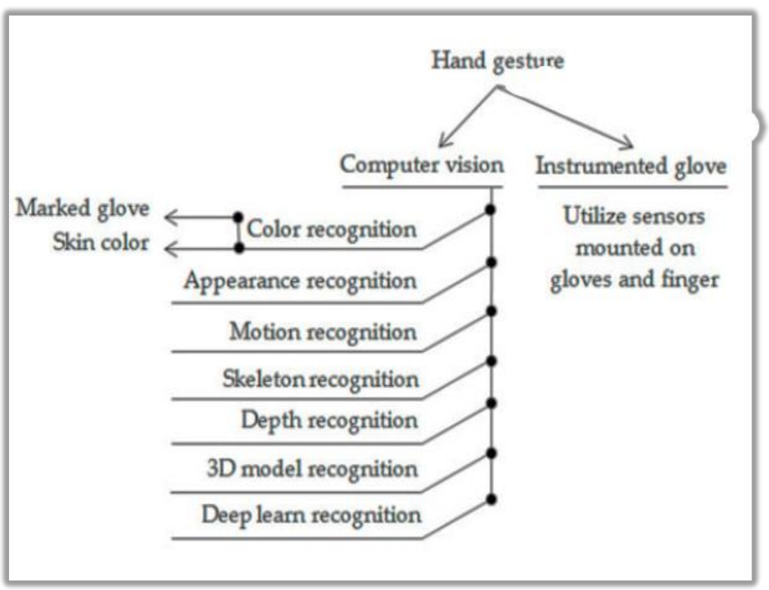


Figure 2.5: Computer Vision for Hand Gesture Recognition [2].

2.3.3 Sign Language Recognition

Sign language is widely regarded as the primary means by which the deaf and dumb can communicate with the hearing, but most hearing people know little to nothing about sign language. That's why it's important to find new methods of reaching out to them. strategies that can make life easier for the hearing-impaired and the speech-impaired. a plethora of uses The deaf community can benefit greatly from deaf service. Images and symbols in sign language have a unique property that allows them to translate between spoken language and the language of the deaf.

Hand signs, face emotions, and other body language are the basis of sign languages [37]. Learning sign language is beneficial for people of all ages and backgrounds, not just those who are deaf or hard of hearing, who use it as their primary way of communication. Others who may benefit from learning sign language include those with developmental disorders such as Autism, Apraxia of Speech, Cerebral Palsy, and Down Syndrome.

There is no universally accepted sign language. There are as many diverse sign languages as there are spoken languages because each has evolved independently as different groups of people have interacted with one another. The number of sign languages in use now is estimated to be between 138 and 300 [38].

It's interesting to note that even when a country and another share a spoken language, they don't always use the same sign language. American Sign Language (ASL), British Sign Language (BSL), and Australian Sign Language are all variations of the English language [5].

Most beginning sign language students begin their studies by learning the alphabet. 'Fingerspelling' refers to the practice of using one's hands to stand in for each letter of the alphabet. Used by signers to manually spell out names of persons, places, and things for which there is no standard sign [18], it is an essential tool. Visual cues such as hand gestures, face and body alignment, lip movement, and facial expressions and emotions form the basis of sign language. Arabic Sign Language (ArSL), American Sign Language (ASL), and Indian Sign Language are only a few examples of the various sign languages that exist, each with its own unique features because they are based on different spoken languages and their regional dialects. Isolated sign language refers to communication that consists of individual hand gestures for individual words, while continuous sign language refers to communication that consists of sequences of gestures that create sentences. Most sign language categorization systems rely heavily on the identification of hand gestures [39].

Hand gesture detection has recently increased by research on automated [1] due to several factors, including the growing population of the deaf and hard of hearing, the popularity of gesture-controlled smart gadgets, games, and assistive technology, and the need for communication aids for these populations. Accurate hand gesture recognition can aid in the development of a powerful sign language recognition and classification system, which can enable both hearing- and speech-impaired people to communicate more effectively. Conventional algorithms find it difficult to uncover robust features when temporal misalignment is present, machine learning methods that don't use manually created features can't tell the difference between crucial and irrelevant parts of each frame [40].

2.4 Hand Detection and Segmentation

The accuracy of a gesture recognition algorithm relies heavily on the accuracy of its hand detection and tracking [36]. Hand detection aims to locate the human hand in an image, while hand segmentation seeks to isolate the hand from its background. Segmentation is a common image processing technique. The best method for segmenting hands depends on image quality, the dataset, and static or dynamic hand movements. Image segmentation identifies a region of interest (ROI). It divides an image into sections. Image objects are image fragments. Similarity, discontinuity, etc. are considered. Segmenting images simplifies the image for better analysis.

2.4.1 Hand Detection-Based Machine Learning

It is common to use a K-means clustering to sort a dataset if the labels are obscure. In this case, the purpose is to identify specific groups based on the number of groups represented by K [41].

The FCM formula clusters related image points. Randomly and iteratively assigning cluster coefficients to each pixel achieves this[42]. This algorithm differs from the previous one by allowing one pixel to belong to more than one cluster.

KNN is a simple, easy-to-understand, adaptable machine learning method. KNN is utilized in image processing. The feature-similarity-based KNN algorithm [35]. The concept of similarity is captured by KNN (sometimes called distance, proximity, cosine similarity, or closeness). The Euclidean distance is a popular and familiar choice [43].

In Otsu's segmentation, the input image is first processed and converted to a gray image because threshold Otsu works only on 2D images, and then the histogram is obtained to display the pixel distribution. The peak value is emphasized. Next, compare the threshold value to the image pixels. If the pixel is above the threshold, set it to white. The thresholds in Otsu's are set automatically [44].

The HSV color spaces for hand detection, HUE, consist of the primary colors of red, blue, and yellow and their complementary secondary colors of orange (green) and violet (violet) on a color wheel or circle (Figure 2.6). What you're referring to when you use the term ‘hue’ is the pure color or the rainbow's visible range of primary colors [45].

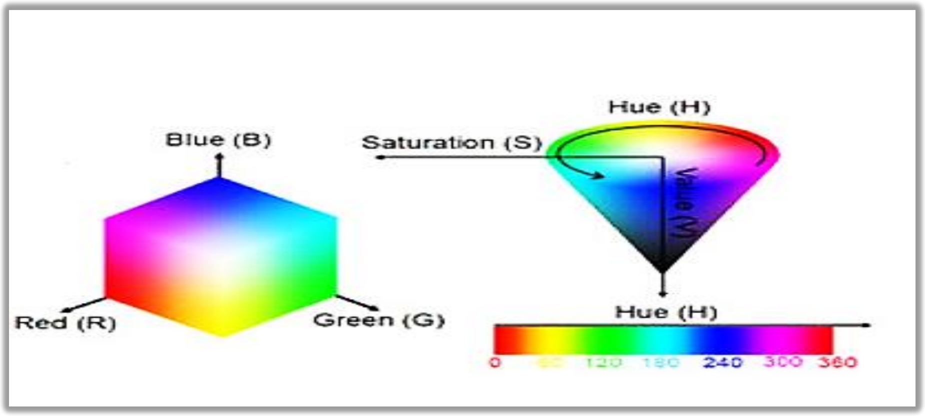


Figure 2.6: RGB vs. HSV Color Space [45].

Many algorithms find image edges used to detect the object, like the Canny algorithm. Canny describes three criteria to evaluate an edge detection algorithm [46] Good detection, there should have a low likelihood of missing true edge points and mistakenly labeling non-edge locations. Good localization, the operator's edge points should be as close to the edge's center as possible. Only one response per edge: since two responses to the same edge must be false.

In (Figure 2.7) used a variety of machine learning segmentation approaches and image processing methods to detect hands in a series of images with varying illumination conditions in order to find the best results.

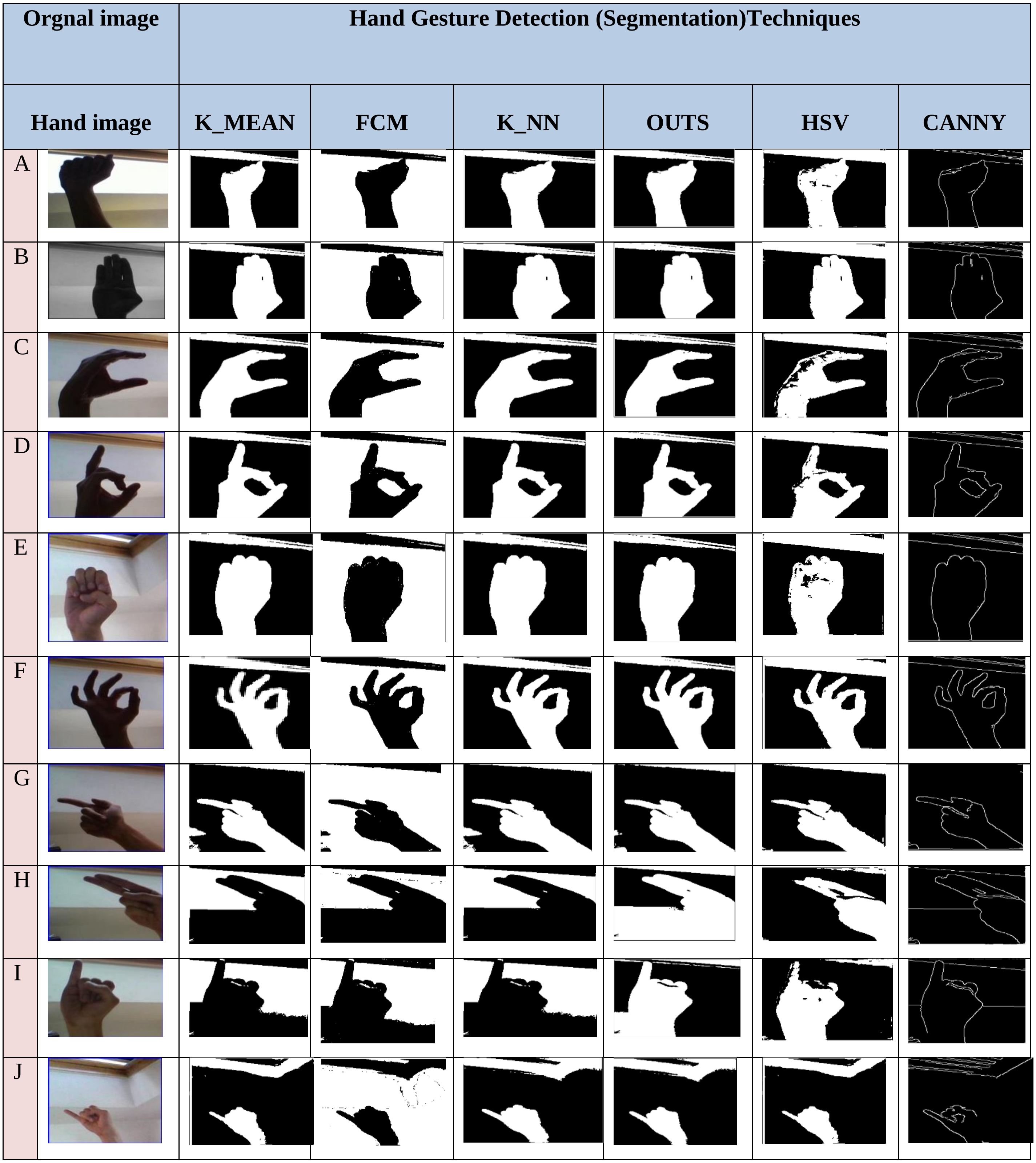


Figure 2.7: Results of Hand Gesture Detection (Segmentation) Techniques [41,42,35,44,45,46]

In (Figure 2.7) shows that clustering works well when the hand's color gradient is far from the gradient of the background, as shown by the clustering results for the letters A, B, C, D, E, F, G, and J. However, when the hand's color gradient is closer to the gradient of the background, the separation process becomes more difficult. It is difficult to distinguish between the hand and the background, as in the letters H and I.

2.4.2 Hand Detection Based Deep Learning

The primary goal of both hand detection and hand segmentation is to determine the location of a human hand in an image. Segmentation is a common image processing technique. The best method for segmenting hands depends on image quality, the dataset, and static or dynamic hand movements. Image segmentation identifies a region of interest (ROI). It divides an image into sections. Image objects are image fragments Similarity, discontinuity, etc. are considered. Segmenting images simplifies the image for better analysis [47].

In dynamic hand motion, deep learning methods are used to estimate the hand from an image or video by calculating the spatial location of the main hand joints, as well as the hand and finger tracking solution [48]. With this technology, would be able to discern the shape, motion, and movements of hands. Hand and finger tracking could be used for a variety of purposes, ranging from gesture recognition to gesture-based control of a human-machine interface. In a VR or AR world, it's critical for communication and interaction as well.

Basically, MediaPipe is a framework for computer vision and deep learning, MediaPipe develops perception pipelines [49]. Just know that perception pipelines are some kinds of audio, video, or time-series data that capture the process in the pipelining zone at this time (Look figure 2.8).

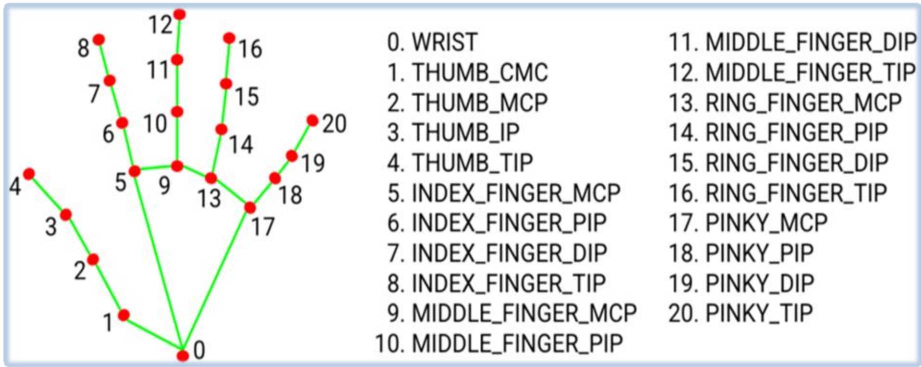


Figure 2.8: 21 Different Landmarks on The Hand [49].

OpenCV can be used to continuously capture images from a camera. Hand tracking uses two modules to determine the hand landmarks points:

Palm detection: Crops the image of the hands to focus on the palms only. Using the BlazePalm detector, (a single-shot) palm detector model, one may infer the palm's bounding box [50]. It saves computing power because it only applies to the first frame. At this stage, the palm is determined by drawing the triangle that indicates the palm of the hand, allowing it to be used in the subsequent hand landmark model for greater accuracy.

Hand Landmarks: Following the initial stage, which involved locating the palm of the hand, the essential points (landmarks), which are the sites of the joints, are now sketched. A circle is painted around each joint that is discovered. The landmark module examines the cropped image and identifies a total of 21 distinct landmarks on the hand. In order to detect landmarks and draw points on the hand, the drawing object (mp. draw) must first be constructed. If the if statement returns true, then the for loop will execute, and a point will be drawn wherever the landmark was found[51]. To create connections between points (lines), connections between landmarks are drawn with the use of a hand object (mp. hand.

HAND CONNECTIONS). The main steps of the hand detection algorithm using MediaPipe are shown in (Figure 2.11)

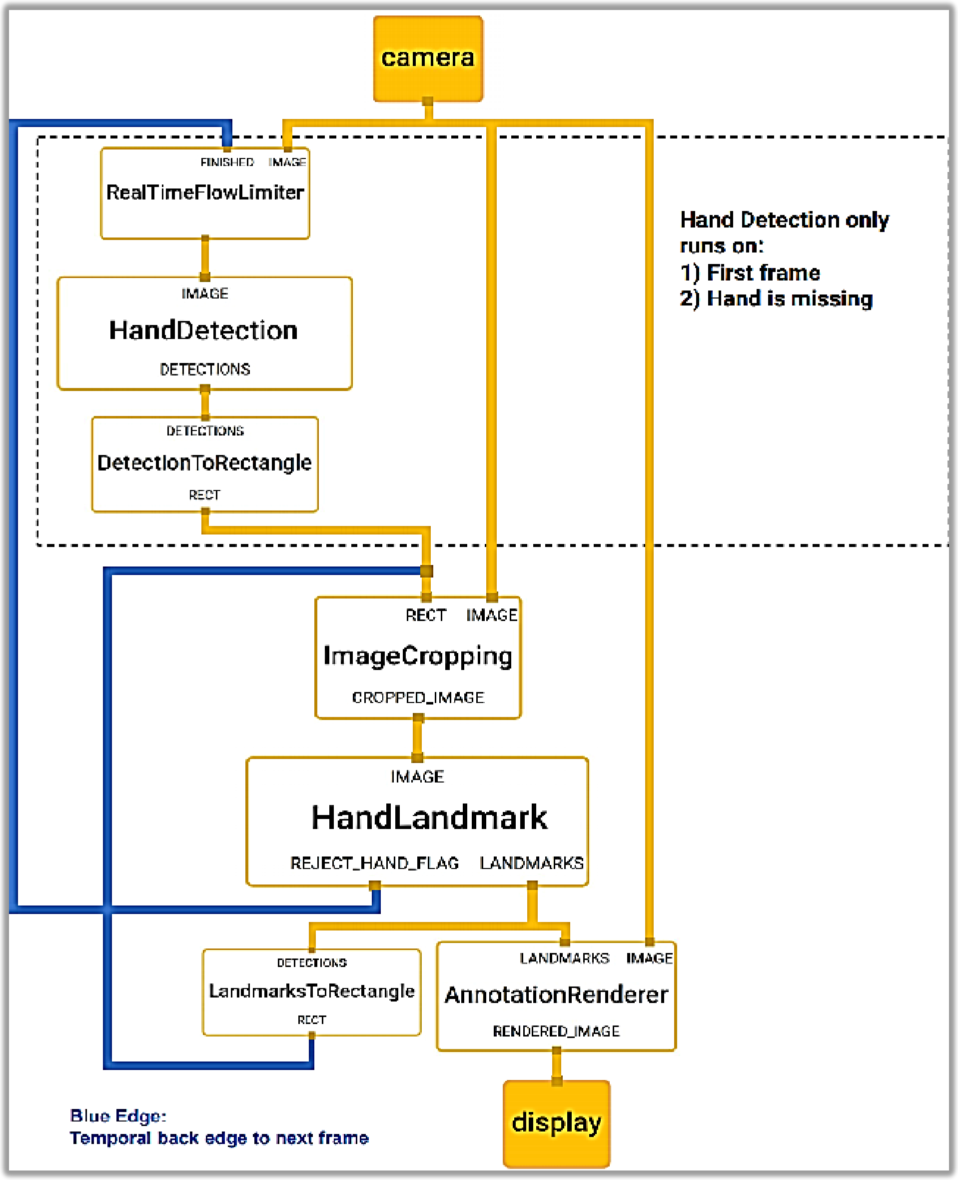


Figure 2.9: Landmark Point for Hand Detection [51].

After that, the hand landmark model would use regression to conduct precise key point localization of the 21 landmark coordinates contained within the cropped hand bounding box. The model is strong even in the presence of self-occlusions and hands that are only partially visible. In the event that the confidence score drops below a predetermined threshold, the palm detection model will be reapplied to the subsequent frame[51].

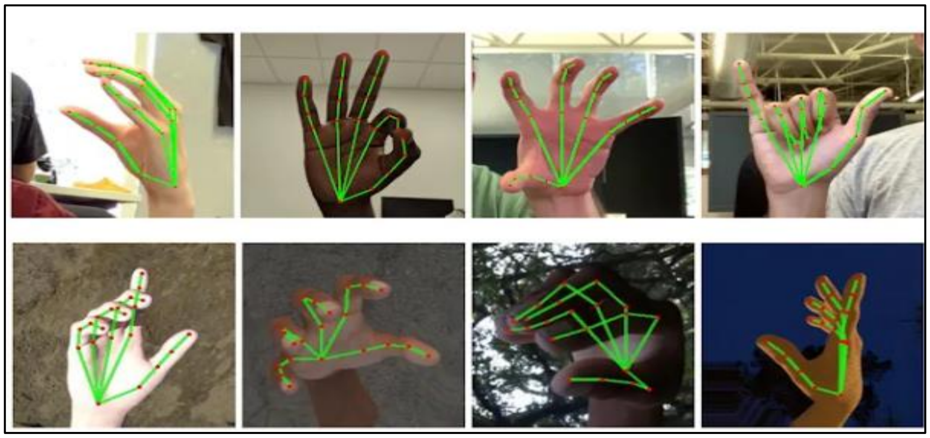


Figure 2.10: Landmarks Hand [50].

Using a deep learning model to estimate the spatial location of important body joints, pose estimation is a computer vision technique for determining a person's pose from an image or video (Figure 2.10). Hand and finger tracking solutions use machine learning to estimate 21 important 3D spatial locations (landmarks) of a hand from a single image frame, just like pose estimation does. It would enable us to recognize hands and hand motions. Hand and finger tracking would open the door to several uses, such as the interpretation of sign language and the manipulation of the human-machine interface by gestural input. It's also crucial since it facilitates conversation and interaction in a simulated or augmented reality setting.

2.5 Associative Memory

Neural Network is considered one of the most important areas used in artificial intelligence in various fields, and it has dealt with many technical problems that require thinking comparable to human thinking. One of the important intelligent artificial neural networks widely used in pattern recognition is the associative network[52].

Associative memory is an essential brain function. learn to immediately recognize a person upon sight, even if they are wearing sunglasses or look far older than remember them being. In order to link vectors (input and output patterns) together, associative neural networks are employed. In associative memory, the input patterns are fed into a neural network in the hopes that they would trigger the generation of the corresponding output pattern[52]. A powerful associative memory can remember many different patterns. During recall, the memory is stimulated by a key pattern (also known as the search argument) that contains some information about a specific pattern in the memory's pattern set. Associating the key pattern with the memorized information allows for recall of the specific stored prototype.

Parallel searching is typically possible in associative memory. The goal of the search is to return a single result or a collection of results that all share a common value with the provided search argument (Figure 2.11).

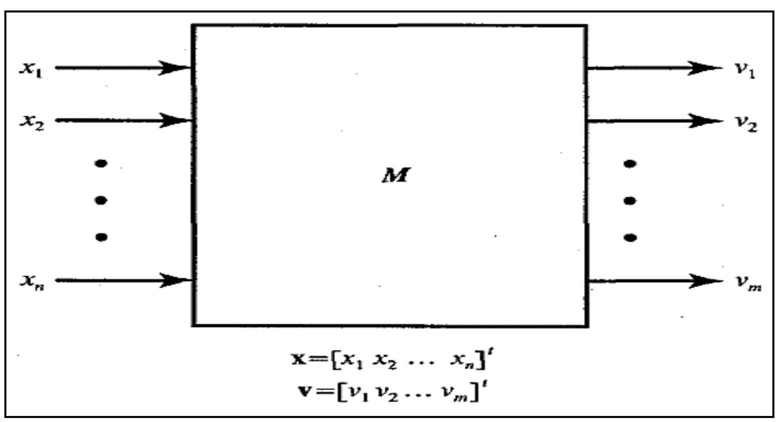


Figure 2.11: A General Block Diagram of Associative Memory [53]

carrying conducting an associative mapping (see Equation 2.1) between a vector x and a vector v.

$$v = M[x] \qquad (2.1)$$

Content-Addressable Memory is another name for these types of storage devices. With the stored patterns as data files, associative memory does a parallel search. The two forms of associative memory are as follows:

1. **Auto Associative Memory:** For the architecture of the Auto Associative memory network, which consists of 'n' input training vectors and an equal number of output target vectors.
2. **Hetero Associative memory:** For the construction of a Hetero Associative Memory network, which takes in 'n' training vectors and outputs 'm' target vectors[54].

3. Multi-Connect Architecture MCA Associative Memory

Standard neural networks have done well at things like processing images, recognizing speech or patterns, and so on [55]. The challenges behind associative memory, such as the capacity for the sadness that you need in memory, as well as the size of the network in most cases, made the researchers not use associative memory with the recognition of hand gestures. The single-layer network of associative memory neural network, which makes it easier to understand and use than other neural networks. MCA It is an update of the Hopfield neural network.

Image processing and speech or pattern recognition are areas where standard neural networks have performed well [55]. The single-layer network of associative memory is useful for the association of patterns. There are fewer moving parts in this neural network, which makes it easier to understand and use than other neural networks.

MCA solved the problem of large storage by building a small network structure that is completely connected with only four weights that will be stored. For each image or video frame, and after getting a binary image of the hand that will only be converted into a vector, the MCA can be applied to a small number of neurons to create a small architecture [56]. In order to speed up MCA computations, rendering, and the real-time process. The number of neurons in the network is limited to three, yet there are numerous connections between these three neurons. Because of this, the smallest network size can be handled by having only four of these connections at all times (Look at Figure 2.12).

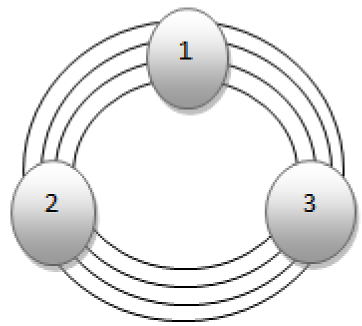
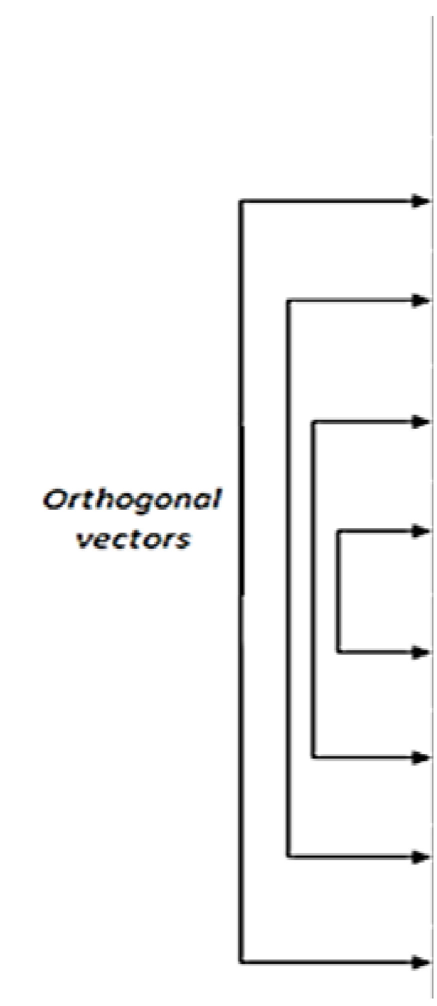


Figure 2.12: The Multi-Connect Architecture (MCA) Associative Memory [56].

The pattern's weight is represented by the connections between these nodes. Only four weights are needed because the maximum number of vectors is 2^3 [56]. Since half of these vectors are orthogonal (see figure 2.13) to the other half, and they have the same weight matrix, this is the reason why half of these vectors are orthogonal. The MCA algorithm has two phases the learning and convergence phases.

It is worth noting that obtained the weight matrices by multiplying the vector by itself and making the main diagonal zero. The significance of the MCA is the fact that multiple connections are required to express the weights of these vectors. Since each vector has three elements, the number of nodes is equal to the number of those elements, and the number of connections is equal to how many points there are in each vector, which equals a total of four.



Binary representation	Bipolar vector	Weights for swv / twv	**majority description** smd
0 0 0	-1 -1 -1	$\begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$	-1
0 0 1	-1 -1 1	$\begin{bmatrix} 0 & -1 & -1 \\ -1 & 0 & 1 \\ -1 & 1 & 0 \end{bmatrix}$	-1
0 1 0	-1 1 -1	$\begin{bmatrix} 0 & -1 & 1 \\ -1 & 0 & -1 \\ 1 & -1 & 0 \end{bmatrix}$	-1
0 1 1	-1 1 1	$\begin{bmatrix} 0 & 1 & -1 \\ 1 & 0 & -1 \\ -1 & -1 & 0 \end{bmatrix}$	1
1 0 0	1 -1 -1	$\begin{bmatrix} 0 & 1 & -1 \\ 1 & 0 & -1 \\ -1 & -1 & 0 \end{bmatrix}$	-1
1 0 1	1 -1 1	$\begin{bmatrix} 0 & -1 & 1 \\ -1 & 0 & -1 \\ 1 & -1 & 0 \end{bmatrix}$	1
1 1 0	1 1 -1	$\begin{bmatrix} 0 & -1 & -1 \\ -1 & 0 & 1 \\ -1 & 1 & 0 \end{bmatrix}$	1
1 1 1	1 1 1	$\begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$	1

Figure 2.13: All the Possible Vectors with Their Weight Matrices and Majority Description [56].

It is worth mentioning that the majority description is the sign resulting from the sum of the three elements of the vector, each with -1, whose origin is 0.

During the training procedure depicted in Algorithm 2.1[53], these weights do not necessitate any computation.

Algorithm 2.1: MCA Learning Phase.
Input: training patterns *p*. Output: lookup table for all *n* corresponding stored patterns.
Step 1: Initialize the four connection weights matrices. $w_0 = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}, \quad w_1 = \begin{bmatrix} 0 & -1 & -1 \\ -1 & 0 & 1 \\ -1 & 1 & 0 \end{bmatrix}, \quad w_2 = \begin{bmatrix} 0 & -1 & 1 \\ -1 & 0 & -1 \\ 1 & -1 & 0 \end{bmatrix}, \quad w_3 = \begin{bmatrix} 0 & 1 & -1 \\ 1 & 0 & -1 \\ -1 & -1 & 0 \end{bmatrix}$ Step2: Repeat step 2.1 to the end of training pattern *p*: Step 2.1: Divide the training pattern *p* to *n* vectors *v* with length three. Step 2.2: For each vector *v*, repeat steps 2.2.1, 2.2.2 and 2.2.3: Step 2.2.1: Assign the vector majority description *smd* as follows: $md(v) = \sum_{i=1}^{3} v_i$ $smd = \text{hard} - \text{limiter}(md(v)) \begin{cases} 1 & md(v) \geq 1 \\ -1 & md(v) \leq 0 \end{cases}$ Step 2.2.2: Assign the stored vector's weight *svw* as follow: $svw = f(Dcode(v)) \begin{cases} 0 \;\; or \;\; 7 & 0 \quad \{means\, w0 \\ 1 \;\; or \;\; 6 & 1 \quad \{means\, w1 \\ 2 \;\; or \;\; 5 & 2 \quad \{means\, w2 \\ 3 \;\; or \;\; 4 & 3 \quad \{means\, w3 \end{cases}$ where: Dcode is a function to convert the binary number to decimal number. Step 2.2.3: Save *smd* and *svw* for this vector in the lookup table. Step 3: End.

The convergence process repeats the same operations that were performed in the learning phase, but by adding the energy function on which the algorithm relies in determining the extent of convergence. The steps of the convergence process are shown in the Algorithm 2.2[53].

Algorithm 2.2: MCA Convergence Phase.

Input: *n* of unknown patterns *p*.
Output: Convergence pattern *CP*

Step 1: Initialize the four connection weights matrices.

$$w_0 = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}, \quad w_1 = \begin{bmatrix} 0 & -1 & -1 \\ -1 & 0 & 1 \\ -1 & 1 & 0 \end{bmatrix}, \quad w2 = \begin{bmatrix} 0 & -1 & 1 \\ -1 & 0 & -1 \\ 1 & -1 & 0 \end{bmatrix}, \quad w_3 = \begin{bmatrix} 0 & 1 & -1 \\ 1 & 0 & -1 \\ -1 & -1 & 0 \end{bmatrix}$$

Step 2: Initialize the energy function matrix *e*:

$$e = \begin{bmatrix} -3 & 1 & 1 & 1 \\ 1 & -3 & 1 & 1 \\ 1 & 1 & -3 & 1 \\ 1 & 1 & 1 & -3 \end{bmatrix}$$

Step 3: Repeat steps 3.1, 3.2 and 3.3 until the unknown pattern *p* is ended:

Step 3.1: Divide the unknown pattern *p* to *n* vectors *v* with length three.

Step 3.2: Assign the test vector's weight *tvw* for all vectors *v* of the test pattern as follows:

$$tvw = f(Dcode(v)) \begin{cases} 0 \quad or \quad 7 & 0 \quad \{means\, w0 \\ 1 \quad or \quad 6 & 1 \quad \{means\, w1 \\ 2 \quad or \quad 5 & 2 \quad \{means\, w2 \\ 3 \quad or \quad 4 & 3 \quad \{means\, w3 \end{cases}$$

Step 3.3: Sum up the energy function for all *n* vectors in the unknown pattern each with its corresponding vector in the stored patterns:

$$ep = \sum_{i=1}^{n} e[svw_i, tvw_i]$$

Step 4: Determine the stored pattern number *minp* with the minimum energy function to converge the unknown pattern towards it:

$$minp = \min(ep)$$

where the min function is to determine the minimum energy function in *ep* array.

Step 5: Repeat steps 5.1, 5.2 and 5.3 to build the final converge pattern *cp:*

Step 5.1: Assign *temcv* for each *n* vector in the test unknown pattern

$$tempcv_i = v_i \times svw_{minp}$$

Step 5.2: Assign the result vector majority description *rmd* to each *n tempcv* vector.

$$md(tempcv^i) = \sum_{j=1}^{n} tempcv_j^i$$

$$rmd = \text{hard} - \text{limiter}(md(tempcv_i)) \begin{cases} 1 & md(tempcv_i) \geq 1 \\ -1 & md(tempcv_i) \leq 0 \end{cases}$$

To match the unknown pattern with the known patterns in the lookup table (i.e., the weight and sign), an energy function (e) is used. When comparing the vectors in the unknown pattern to those in the lookup table, the energy function's sum is calculated; the value of this function is dependent on the associated weight of each vector. Potential energy functions can be calculated and assigned in a two-dimensional matrix (see Table 3) because vectors and weights are constant [72]

$$E = -\frac{1}{2}\left(\sum_{i=1}^{n}\sum^{n} w_{ij} v_i v_j + \sum_{k=1}^{n} t_k v_k\right) \qquad (2.2)$$

Where:

N are elements in the vector ***v***.

Wij the weight from neuron ***i***'s input to neuron ***j***'s output is known

t threshold value in this network=0.

From the (Table 2.1) all the similar weights get -3. As for the non-similar, you get a value of 1. This helps in giving the lowest value the most similar and identical patterns. Calculating the decimal number for each vector will determine the weights.

Table 2.1: Values of the Energy Function that are Related to Each Vector.

State no.	7 or 0	6 or 1	5 or 2	4 or 3
7 or 0	-3	1	1	1
6 or 1	1	-3	1	1
5 or 2	1	1	-3	1
4 or 1	1	1	1	-3

The final stages after the end of the convergence process are important in reconstructing the pattern with which the convergence occurred based on the stored weights.

$$CVi = vi \times svw\ minp \qquad (2.3)$$

Hear assign *cv* for each n vector *and* create the converge vector *CVj* in the converge pattern *cp*

$$CV_j = (smd_{\ minp} \times tmd)\ cv_i \qquad (2.4)$$

2.6 Sign Language Recognition System Evaluation

Because the accuracy, percentage, and other criteria for evaluation are dependent on the findings of the confusion matrix, it is commonly utilized in the process of evaluating systems. In contrast to binary systems, this one does not have categories of either 1 or -1. Because there are no positive or negative classes, determining TP, TN, FP, and FN may at first appear to be a difficult task; nevertheless, in reality, the process is relatively straightforward [57]. In this part of the process, have to figure out the TP, TN, FP, and FN values for each class individually, as illustrated in (Figure 2.14).

Instances that can be classified as positive with absolute certainty are referred to as True Positives (TP). A positive prediction that turns out to be incorrect, often known as a false positive (FP), refers to an instance that ends up belonging to a different, negative class. "False Negatives" (FN) are situations that people mistakenly believe are in the negative category when, in reality, they belong in the positive class. The members of the class who correctly predicted that the class would be negative make up the negative class.

Negative classes are those that have been successfully predicted as True Negative (TN).

Actual Class

Predict Class	A	B	C
A	10	2	3
B	4	9	2
C	3	1	11

TP for A =10

FP for A = (2+3) = 5

FN for A = (4+3) = 7

TN for A = (9+2+1+11) = 23

Figure 2.14: Confusion Matrix for Multiclass [57].

Correctly categorized examples (TP+TN) as a percentage of all examples (TP+TN+FP+FN) is a common way to express an algorithm's accuracy.

$$\text{Accuracy} = \frac{TP+TN}{(TP+TN+FP+FN)} \tag{2.5}$$

Precision is the ratio between the True Positives and all the Positives.

$$\textbf{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{2.6}$$

The recall is the measure of the model correctly identifying True Positives.

$$\textbf{Recall} = \frac{TP}{TP + FN} \qquad (2.7)$$

2.7 Summary

This This chapter reviewed the literature related to the concept of hand gesture recognition, and the chapter structure was presented in several sections. The first section dealt with gesture recognition methods, which fall into two different categories. the second section show applications of the gesture recognition system. The third section is concerned with detecting hand gestures. A number of machine learning methods were viewed within the gesture detection stage, and then the hand detection method was presented using the landmark point. In the fourth part of the research, the idea of associative memory and its different types are reviewed. The MCA method, which was used to recognize hand gestures in this thesis, is also explained. The confusion matrix for multiclass was also offered in the final section as a means of rating the most popular hand gesture detection system.

CHAPTER THREE

REAL -TIME HAND GESTURE RECOGNITION

3.1 Introduction

This chapter details the process followed to develop an effective RTHGR system, which facilitates interaction between sign language users and non-users by translating their language into text. The proposed system consists of three basic components, which start with detecting the human hand, then recognizing the gesture of this hand, and finally translating this gesture into text. The overall framework of the RTHGR system with its components is shown in (Figure 3.1).

3.2 Research Methodology

Based on (Figure 3.1) can deduce this proposed system technique. Sign language data will be gathered from video stream frames of various lighting conditions and hand sizes for this illustration. Python was used for the RTHGR's development. The latter is training with images that are apt representations of the letter or number being gestured. Then extract the shape skeleton of the hand by using a landmark point. And then cropping the skeleton of the hand and turning it into a black image that is 200 x 200. Then used MCA to recognize hand gestures. The frames are processed, and the key frame that reflects the sign is recognized, which has a high confidence (convergence rate). Then a quantitative comparison will be made between the proposed system and other works. Display sentence text resulting from recognizing hand gestures, and display the sign as a sentence. In the end in chapter four, Analyze, discuss, and summarize the RTHGR comparative results.

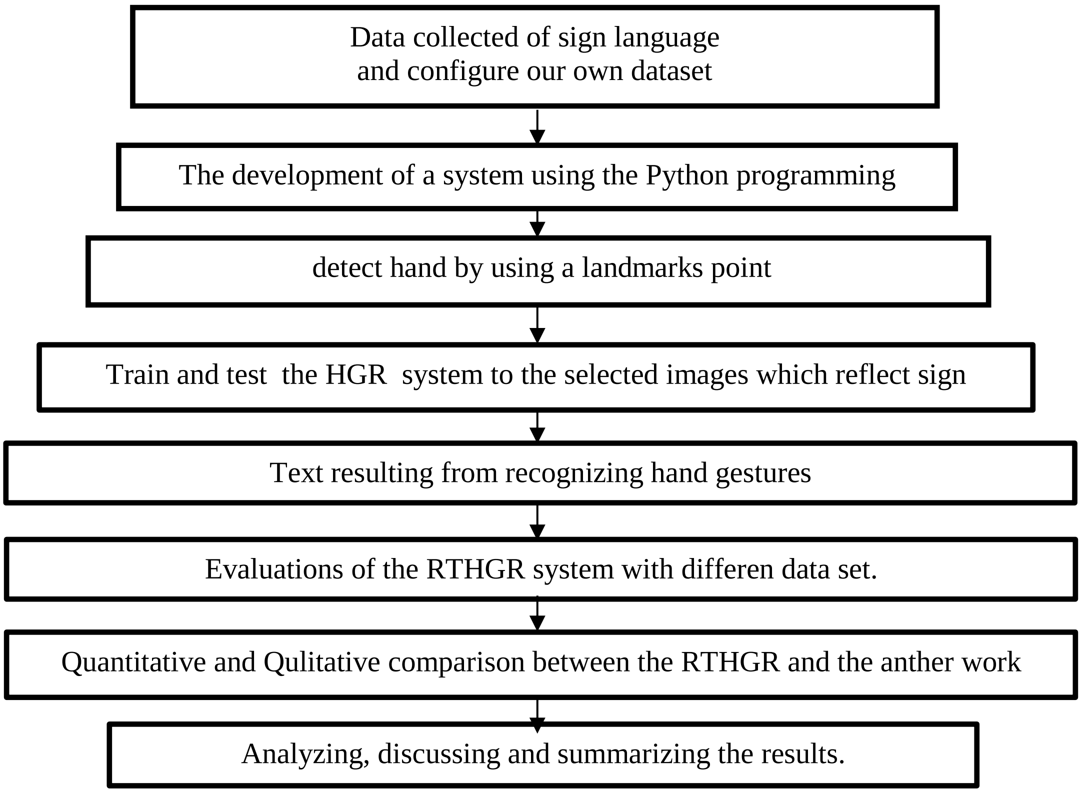


Figure 3.1: Proposed System Methodology Flow

The aim of the RTHGR is to use associative memory to recognize hand gestures in real-time. The main research question is: "What is the best and most efficient technique for recognizing hand gestures in real-time as a new trend? To answer this question, a machine learning approach has been applied in the hand detect and segmentation stage using two modules to determine the hand landmark points. The method of detecting the shape of the hand gave the best results.

In addition to the new method, used Multi-Connect Architecture (MCA) associative memory for classifying the hand gesture. As an interesting trend, associative memory was used for the first time with the recognition of hand gestures. The researchers have avoided using it in this field due to its known limitations in storing the patterns in the memory. In addition to the problem of similarity between patterns, it Doesn't work in real time, most of which are overcome in MCA associative memory. The Proposed system's goal is to develop an effective system for recognizing and correctly interpreting hand gestures using computer vision technology.

The system depends on the appropriate processing and analysis of the video stream of hand gestures taken from the laptop camera, so in fact, need to convert the video into images, treat each frame as an image, process it, move to the next frame, and so on. The system was applied to the data set (American Sign Language), which is a sequential frame for the sign of the hand. As for the Arabic and Chinese sign languages, they were configured.

The stage of detecting the hand gesture is one of the important stages on which the next step greatly depends. Many artificial intelligence systems have been used, some of them are good, and some of them should be in a controlled environment. The problem of illumination and background clutter is one of the most important problems that researchers face in the gesture detection phase, as shown in Chapter Two, Figure 2.7. The stage of recognizing hand gestures is a very important one.

3.3 The Proposed System

The concept of associative memory was used as a new direction that researchers had not addressed before within the recognition of hand gestures or sign language, where MCA was used in the stage of recognizing sign language. The use of MCA gives good results with an efficient proposed system.

Making the system applicable and technologically important because it is considered an important tool to make communication between deaf people and ordinary people more flexible is an important part of the proposed system. In the stage of displaying the gestures flowing from the video in the form of text, the processing includes taking into account the number of classes in the look-up table and the possibility of giving priority to displaying the results of sign language recognition for the frame that has the highest confidence among several sequential frames.

This Proposed system deals with the image segmentation method and applies it to the American Sign Language and another dataset, and uses the MCA associative memory to recognize hand gestures see the steps of the RTHGR system in (Figure 3.2).

Recognizing and interpreting hand gesture motion requires accurate tracking and detection of the hands. Using hand landmark points in the tracking and segmentation process yielded very important results, and since the associative neural network used feeds from a binary image, it was necessary to take the landmark points of the hand as a binary image and feeds for initializing the MCA network. The result of the first part of the proposed system for hand gesture recognition is depicted in (Figure 3.3), and it shows the main steps to detect and segment the hand and initialize the input to the associative memory neural network used to recognize hand gestures. Details of the steps involved in the RTHGR system are shown in (Figure 3.2) and algorethem 3.1.

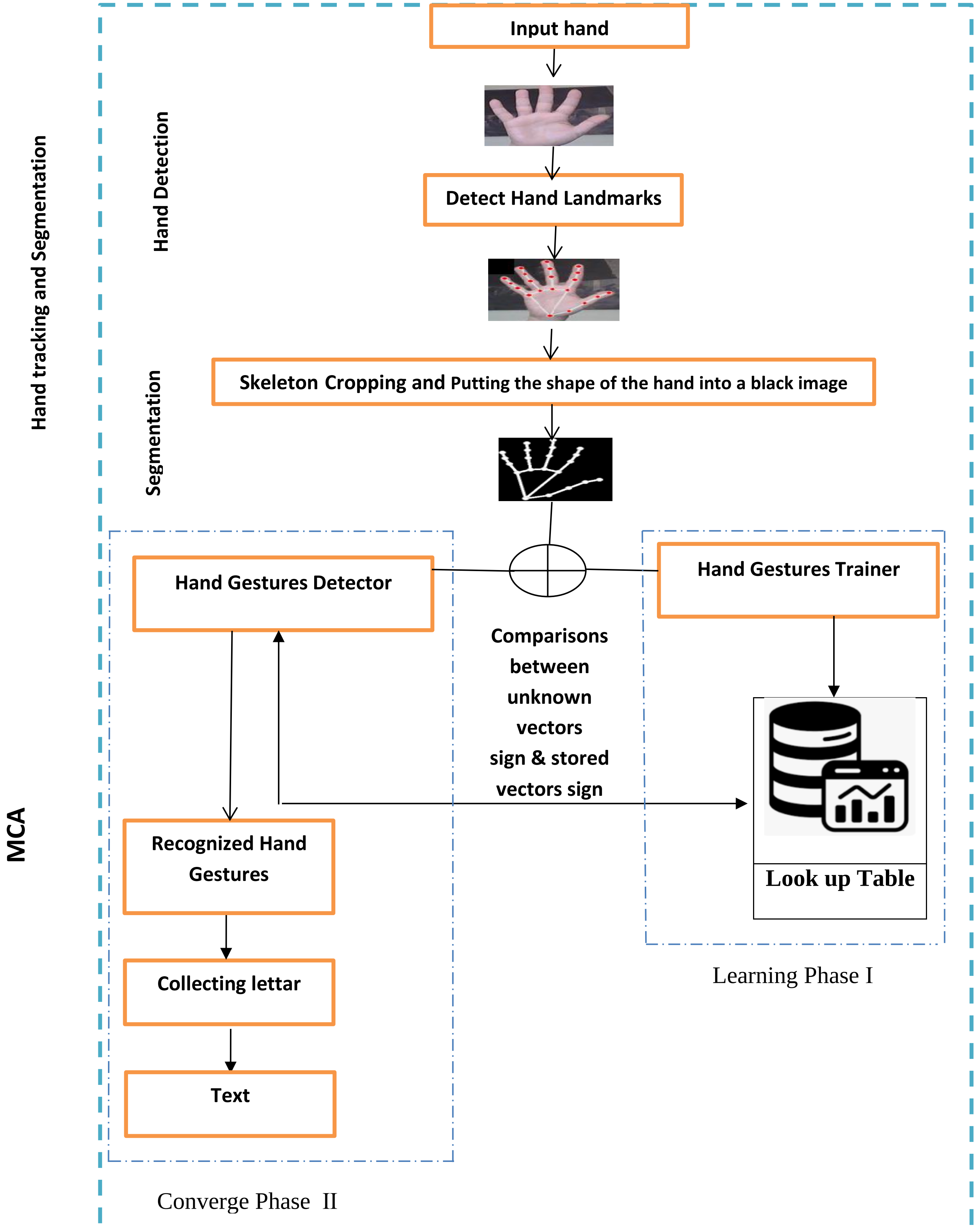


Figure 3.2: Real-Time Hand Gestures Recognition (RTHGR) system.

3.3.1 Hand Detection And Segmentation

The effectiveness of a gesture recognition system is heavily dependent on the accuracy and speed with which it can segment and recognize hands. One objective of hand detection is to locate the exact locatiowen of a hand in an image, while another goal of hand segmentation is to isolate the hand from its surrounding context. In the field of image processing, segmentation is standard and important.

When deciding on the optimal technique for hand segmentation, you should consider image quality, the dataset, and whether hands are static or dynamic. In order to isolate a specific area of interest in an image, segmentation (ROI) is performed. Parts of images are called "objects or hand ". To analyze an image more effectively, effectively segmenting it is the first step.

In dynamic hand motion, deep learning models are used to estimate the hand from an image or video by calculating the spatial location of the main hand joints, as well as the hand, and finger tracking solution. With this technology, would be able to discern the shape, motion, and movements of hands.

The importance of using landmark points lies in giving the features of the skeleton or shape of the hand and, thus, the possibility of obtaining the structure graph of gestures. It has proven its efficiency in detecting the hand.

It constructs perception pipelines just know that perception pipelines are some kinds of audio, video, or time-series data that capture the process in the pipelining zone at this time. In the proposed system, employed the methodology for detecting the hand using the landmark points provided by Google, to segment the hand from the frame as a whole Convolution on the image whole to detect the palm depending on skin color by using (mp.solutions.hands).

Draw the triangle on the palm of the hand, then store the index of the three vertices of the triangle. The landmark points methodology was often used to extract features and take advantage of the coordinates of the landmark points in that. As for the proposed system, used only the landmark points and took only the shape of the skeleton of the hand resulting in the interdependence of the 21 points with each other, and then it is displayed thickly. Can determine the possibility of changing the size of the circles and lines that connect one joint to another. The thickness of the line and circle for each identifier is set in white, as well as the thickness is set to 12 pixels, and the radius of the circle is set to 3 pixels. These standards were taken because they clearly gave shape to the hand structure.

After displaying the skeleton of the fingers of the hand within a specified width and radius of a circle, a drawing of the template is determined on the borders of the skeleton, and the hand is then cropped out of the overall image.

It is worth mentioning here that the drawing of the window on the hand was based on the center, which is the zero point, since the zero point in the palm has been considered the point from which the borders of the templet or window start and it must contain all the points of the landmark. Drawing a hand template to prepare it for the cropping process and to give flexibility in taking the structure in a flexible manner. An additional area of 15 pixels was specified, starting from the starting point (x_min - 15, y_min - 15) and the endpoint (x_max + 15, y_max + 15).

This method has proven to be very effective for detecting any hand and drawing a square or rectangle in a variety of ways, depending on the condition of the hand closed or open. Then put the white skeleton of the hand in the black background image of size 200 * 200. Algorithm 3.1 clarifies hand detection and segmentation.

Algorithm 3.1: Hand Detection And Segmentation

Input: Stream of Frames
Output: Binary Image (Shape of Hand)

Step 1: Read the frames - jpg image.

Step 2: Create a 200*200-pixel black image.

Step 3: BGR is converted to RGB

Step 4:Convolution on the image whole to detect the palm depending on skin color by using Package mediapipe.python.solutions (**mp.solutions.hands).**

Step 5: Draw the triangle on the palm of the hand, then store the index of the three vertices of the triangle by using **mp.solutions.drawing_utils**.

Step 6: Save the id with coordinates landmark point LM. Draw the landmark point for the finger of the hand. using **mpDraw Draw - landmarks (image, handLms, mpHands.HAND_CONNECTIONS).**

Step 7: Determining the thickness of the line and circle for each id: color = (255, 255, 255), thickness = 12 pixels, circle radius = 3.

Step 8: For hand LMs, let the highest and lowest points determine which points of the skeleton of the hand are from (tip, pip ,dip ,mcp) . **When LMs is 21 points of skeleton hand**

Step 9: Drawing the template for hand, Start-point = (x_min - 15, y_min - 15) and

End-point = (x_max + 15, y_max + 15).

Step 10: Cropped image in the template and set it (the shape of the landmark) to a black image have been created in step 2.

Step 11: Return binary image**.**

Step 12: End.

The hand landmark model is strong even in the presence of self-occlusions and hands that are only partially visible. In the event that the confidence score drops below a predetermined threshold, the palm detection model will be reapplied to the subsequent frame see (Figure. 3.3) to illustrate the result of hand segmentation.

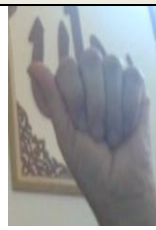
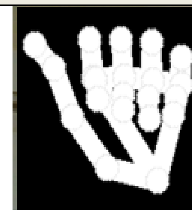
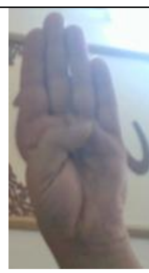
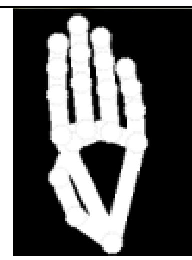
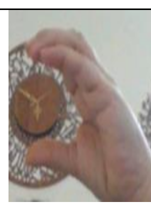
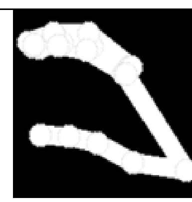
No	Input Hand	Detected Hand	Segmented Hand
1			
2			
3			

Figure 3.3: The sample of hand detection operation.

The landmark points gave an impressive result, especially after overcoming the problems of changing illumination and background clutter, as well as scaling.

Moving on to the training phase, where the MCA network will be fed the binary image produced in the first part (bipolar conversion), As the proposed RTHGR system makes use of MCA, the images that have been filtered will need to be transformed into binary images with a bipolar representation of pixel values

(1 or -1 for each pixel) in order to function properly. This will lessen the total amount of data processing required without compromising any of the image's crucial details.

3.3.2. Recognition of Hand Gestures

Image processing and speech or pattern recognition are areas where standard neural networks have performed well. A single-layer associative memory network is useful for associating patterns of hand gestures. There are fewer moving parts in this neural network, which makes it easier to understand and use compared to other neural networks.

As a result, this Proposed system focuses on to use (MCA), which solves the problem of taking all the similar and dissimilar patterns and focusing on the different patterns so that the pattern is stored for one time only by building a fully connected small network structure with only four weights that will be stored as described in the second chapter. for each image or video frame. So after obtaining a binary image sign pattern (SP) of the hand from the first stage, this image will only be converted into a vector, and the image elements, which are pixel values of either 0 or 255, are captured into a small vector form consisting of three pixels.

The resulting vector pairs contain only the three different elements, and since the length of the vector is 3, the probabilities of the vector elements range from 0 to 7. As shown in the (Figure 2.13). The RTHGR applied to a small number of neurons create a small architecture in order to speed up HGR computations, recognition, and the real-time recognition process. The number of neurons in the network is limited to three, yet there are numerous connections between these three neurons. Due to this, having only four of these connections active at any given time can handle the smallest network size and speed up RTHGR .

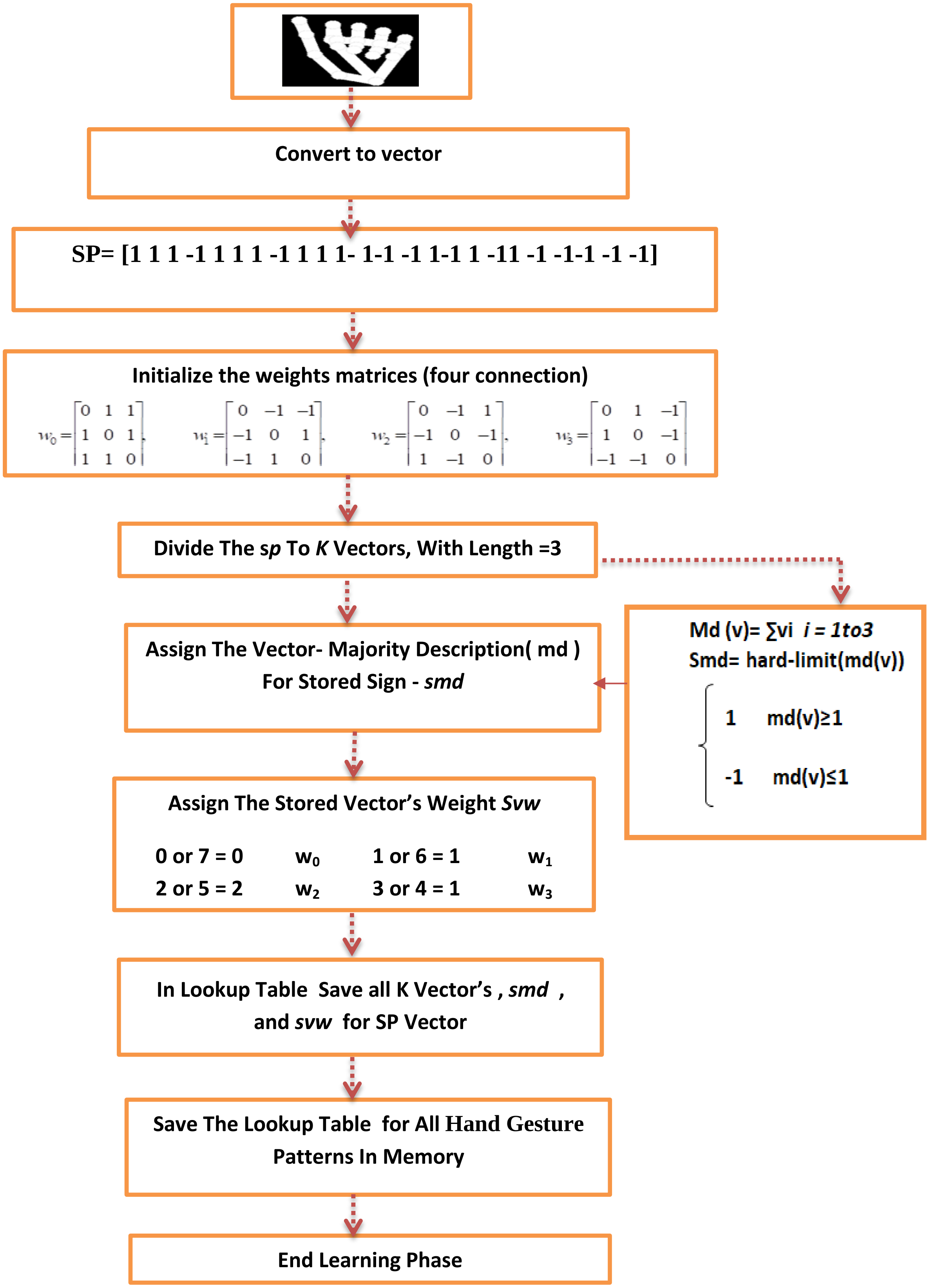


Figure 3.4 : The RTHGR Learning Phase.

Images in the previous step are stored in a 1D array only in case there is no copy of these three elements previously in the new 1D array, and here a 200*200 image will be reduced to a 1D array of 24 elements or less and not more. Potential possibilities to appear in a vector are as follows: [1 1 1], [-1 1 1], [1 -1 1], [1 1- 1], [-1 -1 1], [-1 1 -1], [1 -1 -1], [-1 -1 -1]. Every 0 in the segmented image turned into -1 and every 255 turned into 1.

Algorithm 3.2: Hand Gestures Trainer
Input: Binary Image Output: look-up table
Step 1: Read the binary pattren and convert it into a one-dimensional vector pattern of hand gesture (save only different three vector elements, similar vectors are discarded). **Step 2**: Initialize the four weight matrices. $w_0 = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}, \quad w_1 = \begin{bmatrix} 0 & -1 & -1 \\ -1 & 0 & 1 \\ -1 & 1 & 0 \end{bmatrix}, \quad w_2 = \begin{bmatrix} 0 & -1 & 1 \\ -1 & 0 & -1 \\ 1 & -1 & 0 \end{bmatrix}, \quad w_3 = \begin{bmatrix} 0 & 1 & -1 \\ 1 & 0 & -1 \\ -1 & -1 & 0 \end{bmatrix}$ **Step 3**: The input pattern of hand gesture divide into *K* vectors with length 3. **Step 4**: Assign the vector majority description ***md*** for stored pattern of sign *smd* (1 or -1) md= $\sum_1^3 v$, *when v is vectors of length 3 elements* Smd = hard limit $\longrightarrow$ ***1*** *if md ≥1,* **-1** *if md ≤1* **Step 5**: Compute each vector in decimal, assign the stored vector's weight svw of sign. ***0 / 7 means*** w_0, ***1 / 6 means*** w_1, ***2 / 5 means*** w_2, ***3 / 4 means*** w_3 **Step 6**: Save the result of steps 4, and 5 for this ***K*** vector, ***when K is 8 vectors*** **Step 7**:Save the output of *8 vectors*, *md* and *svw* for each input in *a look-up table*. **Step 8**: End.

The output from the training phase will be a table containing the information that has been calculated. The contents of the table are shown in the (Figure 3.5).

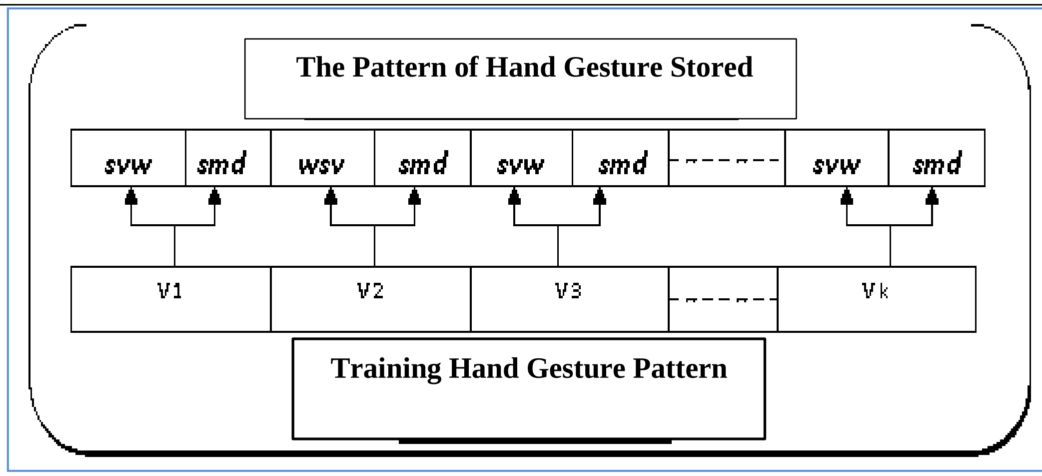


Image Id	Vectors	SVM	SVW
A	[-1, -1, -1]	-1	W0
A	[-1, 1, 1]	1	W3
A	[1, -1, -1]	-1	W3
A	[-1, 1, -1]	-1	W2
A	[-1, -1, 1]	1	W1
A	[1, -1, 1]	1	W2
A	[1, 1, -1]	1	W1
A	[1, 1, 1]	1	W0

Figure 3.5: A Lookup Table for The Training Hand Gesture.

As depicted (Figure 2.12), it is impossible for any node to communicate with itself, but all nodes communicate with each other. Only four weights are needed because the maximum number of vectors is 2^3. half of these vectors are orthogonal to each other, they also share the same weight matrix. This is the main reason why the network is so fast at recognizing hand gestures in real time. Figure (3.4) depicts the steps HGR algorithm for learning.

One of the most significant factors that contribute to the overall effectiveness of the network is the learning phase. The results of the learning phase will determine what comes out of In the convergence phase, Figure 3.6 illustrates it. Despite the fact that the modification was implemented during the learning phase, it is vital to have a full adjustment for the convergence phase in order to guarantee that the RTHGR will be effective overall.

The RTHGR presented a solution to the problem of recognizing hand gestures with little complexity because it is from one layer in addition to a few data sets that were used in training compared to deep learning applications, as well as applying them in real-time and giving results between 95.42 and 92.13% by applying them to more than one data set, in addition to the possibility of benefiting from the current Proposed system and making the recognition of sign language at the level of words and sentences, not just numbers and letters.

The proposed system has already been applied to American Sign Language, Chinese Sign Language, and Arabic Sign Language, as well as the numerals 1–10, and it has the potential to be applied to an infinite number of data sets separately.

The dataset of ASL alphabet images was used in the RTHGR system. As for the dataset for the Arabic and Chinese sign languages, as well as the numbers from 0 to 10, it was produced by us.

In order for the proposed RTHGR by using MCA associative memory to work properly, the images are translated to binary format, with pixel values represented as having a smaller dataset to process during the training phase, either 1 or -1. All four weight matrices (w_0, w_1, w_2, w_3) are initialized, and the unknown gesture is split into k vectors of three elements each. the modification process incorporates the growing importance of the energy function in the convergence process.

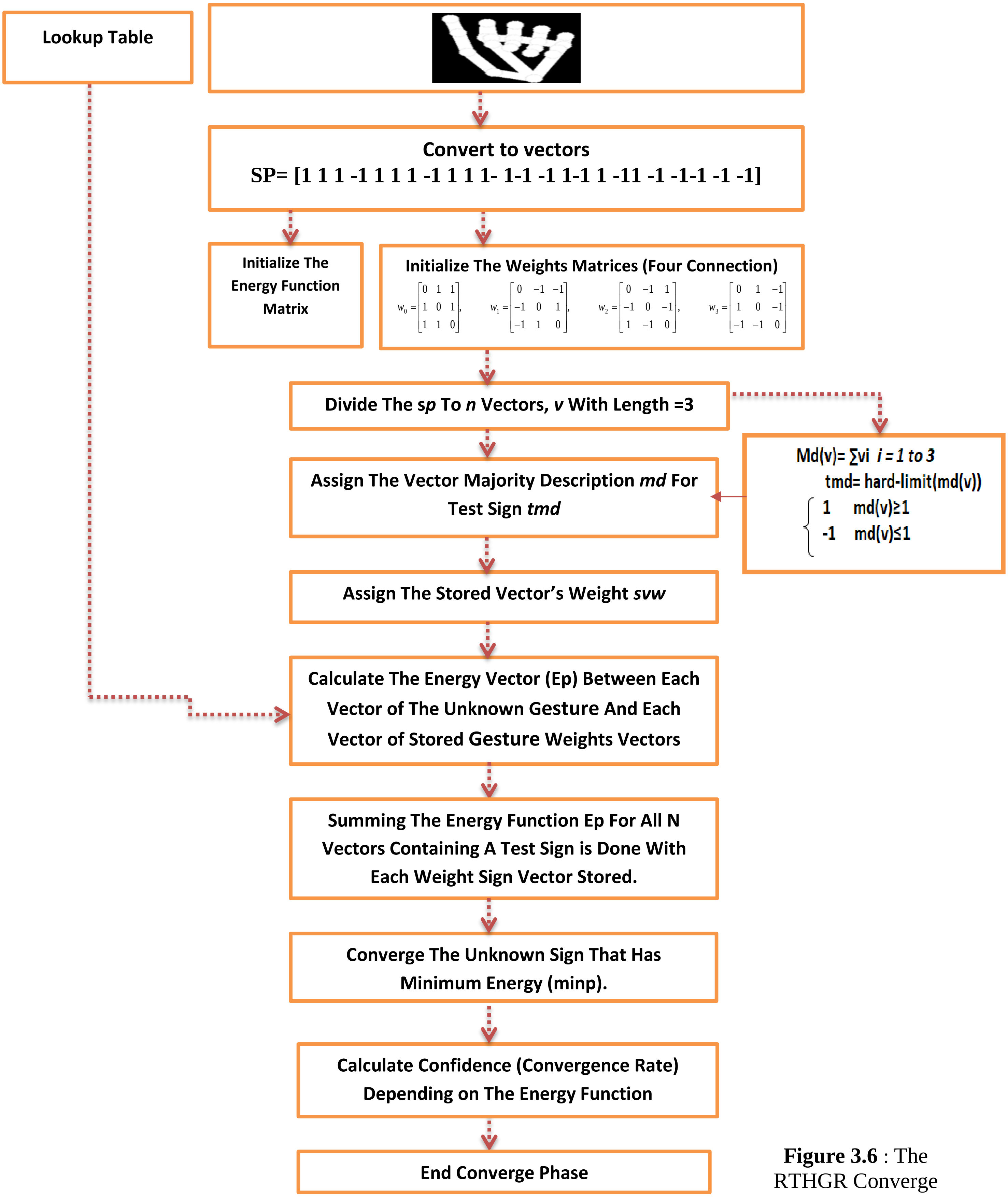


Figure 3.6 : The RTHGR Converge Phases.

After completing the training of a number of images extracting the look-up table and storing it in memory, it is time for the testing process, which depends on the comparison between the stored information in the look-up table as depicted in the Figure (3.5) and the pattern entered in the testing phase**.**

In the convergence phase, the same steps followed in the training phase will be repeated, in addition, the energy function matrix is used to know the extent of similarity between the entered pattern and what is stored. After knowing the closest pattern in the table to the currently tested pattern, calculate the convergence percentage using the following equation.

$$\text{Convergence Rate} = \frac{ep}{The\ Length\ of\ the\ Sign\ pattern} * 100 \quad \ldots\ldots\ldots\ldots \ (3.1)$$

When *(ep)* is Sum up the energy function for all *k* vectors with test hand gesture pattern.

The computation of the energy function explained in chapter two (Table 2.1) in the convergence phase is of great importance because it is responsible for finding conserve the important information of an image the closest gesture pattern stored in the look-up table with the gesture test pattern is a function that reduces steadily to the minimum. when all three-element vectors in the unknown gesture pattern converge to the stored gesture pattern with the lowest energy function, convergence is complete. the mca uses an energy function to find similar gesture patterns to the unknown ones that have previously stored the same information. Algorithm (3.3) explain the steps hand gestures detector.

Algorithm 3.3: Hand Gestures Detector

Input: Binary Image (Test), Look up table

Output: Recognized Hand Gestures

Step 1: Read the binary pattren (test) and convert it into a one-dimensional vector pattern.

Step 2: Initialize the four weight matrices

$$w_0 = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}, \quad w_1 = \begin{bmatrix} 0 & -1 & -1 \\ -1 & 0 & 1 \\ -1 & 1 & 0 \end{bmatrix}, \quad w_2 = \begin{bmatrix} 0 & -1 & 1 \\ -1 & 0 & -1 \\ 1 & -1 & 0 \end{bmatrix}, \quad w_3 = \begin{bmatrix} 0 & 1 & -1 \\ 1 & 0 & -1 \\ -1 & -1 & 0 \end{bmatrix}$$

Step 3: Initialize the energy function matrix.

$$\text{energy function matrix} = \begin{bmatrix} -3 & 1 & 1 & 1 \\ 1 & -3 & 1 & 1 \\ 1 & 1 & -3 & 1 \\ 1 & 1 & 1 & -3 \end{bmatrix}$$

Step 4: Divide the test hand gesture into *k* vectors with length 3.

Step 5: Assign the vector majority description md for tested hand gesture pattern *tmd* .

Md= $\sum_1^3 v$, *when v is vectors of length 3 elements*

tmd = hardlimit $\longrightarrow$ 1 if md $\geq$ 1, -1 if md $\leq$ 1

Step 6: Compute each vector in decimal, assign the tested hand gesture vector's weight *tvw*.

0 / 7 means w_0 , 1 / 6 means w_1, 2 / 5 means w_2 , 3 / 4 means w_3

Step 7: Calculate the energy vector (*ep*) between each vector of the unknown hand gesture pattern and each vector of stored weights hand gesture patterns vectors.

(***ep***) **=** -3 when Wi =Wj , i=j

Step 8: For each stored weights hand gesture patterns vectors, Sum up the energy*(ep)*function for all *n* vectors with test hand gesture pattern.

$$ep = \sum_{i=1}^{n} e[svw_i, tvw_i]$$ **stored vector weight** of the sign , **tested vector weight** of the sign

Step 9: Converge the unknown hand gesture pattern toward the stored weights pattern with minimum energy **(*minp*).**

Step 10: Create the converge hand gesture pattern

Step 11:Depending on the energy function compute the confidence (Convergence Rate) using quasion 3.1.

Step 12: Returns alphabet corresponding to the input sign

Step 13: End.

If the result of the sum of the energy function is -24, then the entered pattern of the gesture or sign is 100% identical with stored sign. If the result of the summation is 8, then the entered pattern of the gesture or signal is not completely identical with stored sign. Figure 3.7 shows an example of calculating the value that will be stored in the lookup table that output of the learning process.

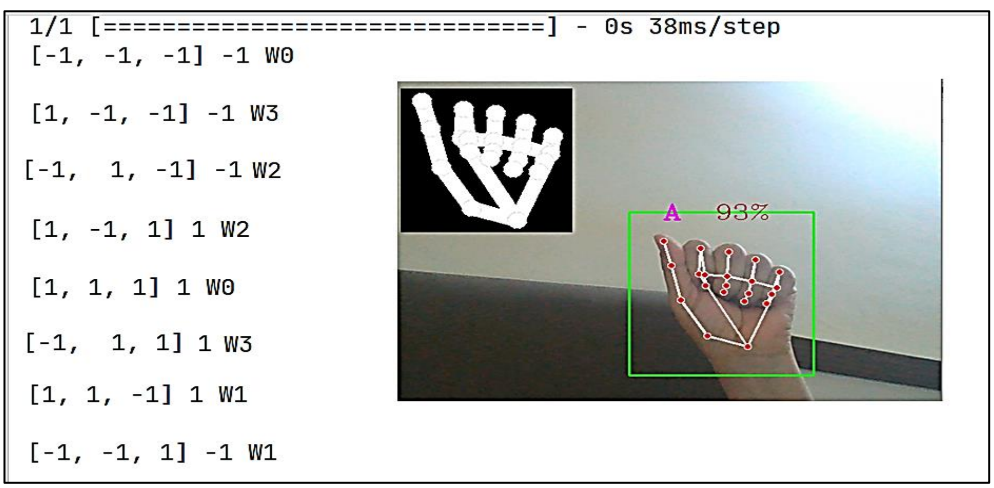


Figure 3.7: Compute The 8 Possible Vectors, *smd* And *weight* For The Pattern.

3.2.3 . Translate Sign Language Into Text

Making the system adaptable and technologically significant since it is regarded as a key instrument for improving the flexibility of communication between deaf individuals and ordinary people is an important component of the proposed system. The processing includes taking into account the number of classes in the look-up table and the possibility of giving priority to displaying the results of sign language recognition for the frame with the highest confidence among several sequential frames at the stage of displaying the gestures flowing from the video in the form of text.

The translation was applied to American Sign Language, which was recognized using the associative memory network as a new direction. In addition, The resulting sentence is more adaptable via the addition of two gestures (delete and space. American Sign Language was learned with the right hand, whereas the digits 1 through 5 and a few other motions (delete and space) were learned with the left. Determining the result of the text and showing the letters as a word depends on the speed and number of frames. Set a time period for the frame rate and choose the frame with the highest confidence to be the result shown on the text box.

Both the right and left hands were used in this system, but the number of classes for the right hand is 28 specified to recognize 28 letters A to Z, in contrast, the number of classes for the left hand is 7, so the speed of the processing of the frames is given to the left hand much less than what is the case for the right hand, Because the number of classes for right hand is more and you need time to search more in the lookup table, and therefore the period granted to the number of frames is greater than the period granted to process the gestures of the left hand.

Algorithm 3.4: Hand Gestures translator **Input:** Alphabet corresponding the input sign **Output:** Text
Step 1: choose No. of frames (period). For both hand **.** Step 2: Take the alphabet of frames with the highest convergence rate Step 3: print the result (letter or alphabet) of sum the frame with the highest convergence rate in the text box Step 4: Collecting the letters resulting from the recognition stage as a word or a sentence from several words Step 5: Return text**.** Step 6: End.

3.4 Summary

The third chapter provided details of the procedures used to achieve the main objective of the Proposed system, which is hand gesture recognition. This goal is to make an efficient system that helps to understand the gestures or sign language used among the deaf and dumb audience. A video-to-image converter and a hand segmentation stage that makes use of landmark points are the first two components that make up the RTHGR, as can be seen in Figure 3.2, a bipolar conversion stage, then a pattern training stage, and a pattern convergence stage, within the MCA tool has been discussion. MCA is a modified associative memory that used as a new direction within associative memory with hand gesture recognition. To apply the system to more than one data set, created a new data set that contains many images of size 200 * 200, especially in Arabic sign language, Chinese sign language, and numbers from 1 to 10. In order to make the proposed system more effective, the signs were shown as text to form sentences consisting of words, not just letters.

CHAPTER FOUR

EXPERIMENTAL RESULTS AND DISCUSSION

4.1 Introduction

This chapter presents the results obtained from the application of the proposed system RTHGR with a discussion and evaluation of the results.

The evaluation is divided into several parts based on the contributions mentioned in the first chapter and evaluated in each part separately. The results obtained from the system will be evaluated in both quantitative and qualitative terms:

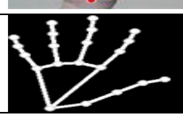
1. Evaluation of the functioning of the proposed RTHGR uses MCA associative memory, to recognize sign language as well as other gestures. in terms of accuracy and the number of training images with efficiency.
2. Real-time Evaluation: it is necessary to evaluate the RTHGR system in real-time.
3. Qualitative evaluation of the RTHGR system process of translation and conversion of signs into words by conducting a questionnaire for several centers and in different governorates. This section presents and discusses the findings from a questionnaire study on the translation of sign language, including details on the study's methodology, community, sample, instrument, validity, reliability, and statistical methods.

The proposed system was implemented in Python language and Supported Platform is PyCharm editor and laptop specifications:

Processor: Intel(R) Core(TM) i5-6600U CPU @ 2.60GHz. **(RAM):** 8.00 GB **Camera Setting**: **Type image**: JPG File **video quality** :720p (16:9, 30 fps). PyCharm 2022.1.1 (Community Edition) **Runtime version**: 11.0.14.1+1-b2043.45 amd64 Windows 10 10.0 Memory: 980M Cores: 4

4.2 Case Study

In this section, an example is taken and the steps of the proposed method are applied to it in sequence from detecting the hand to recognizing the meaning of the input sign show in (Figure 4.1).

1-	Input hand / Unknown sign	
2-	Using landmark point to detect the hand	
3-	Cropping the shape of hand	
4-	Convert the pattern to the vector SP.	SP = [-1 -1 -1 -1 1 1 1 -1 -1-1 1 -1 -1 -1 1 1-1111 -1111]
5-	Divided the sign pattern vector SP for 8 vector that length 3	[-1, -1, -1][-1, 1, 1] [1, -1, -1][-1, 1, -1] [-1, -1, 1][1, -1, 1] [1, 1, -1][1, 1, 1]

6- Assign smd , and tvw for all 8 Vector's. According to the steps (1-7) mentioned in the algorithm (3.3) and calculate the energy vector (Ep) between each vector of the unknown gesture and each vector of stored gesture weights vectors

vector	Decimal	smd	tvw
[-1, -1, -1]	0	-1	W0
[-1, 1, 1]	3	1	W3
[1, -1, -1]	4	-1	W3
[-1, 1, -1]	2	-1	W2
[-1, -1, 1]	1	1	W1
[1, -1, 1]	5	1	W2
[1, 1, -1]	6	1	W1
[1, 1, 1]	7	1	W0

7- Comparisons between test and storage vector in memory tvw,svw.

vector	tvw	SVW	Ep1
[-1 ،-1 ،-1]	W0	W0	-3
[-1 ،1 ،1]	W3	W3	-3
[1 ،-1 ،-1]	W3	W3	-3
[-1 ،1 ،-1]	W2	W2	-3
[-1 ،-1 ،1]	W1	W1	-3
[1 ،-1 ،1]	W2	W2	-3
[1 ،1 ،-1]	W1	W1	-3
[1 ،1 ،1]	W0	W0	-3

vector	tvw	SVW	Ep2
[-1 ،-1 ،-1]	W0	W0	-3
[-1 ،1 ،1]	W3	W2	1
[1 ،-1 ،-1]	W3	W1	1
[-1 ،1 ،-1]	W2	W1	1
[-1 ،-1 ،1]	W1	W0	1
[1 ،-1 ،1]	W2	W3	1
[1 ،1 ،-1]	W1	W2	1
[1 ،1 ،1]	W0	W3	1

Look up table in memory

Image Id	Vectors	SVM	SVW
5	[-1, -1, -1]	-1	W0
5	[-1, 1, 1]	1	W3
5	[1, -1, -1]	-1	W3
5	[-1, 1, -1]	-1	W2
5	[-1, -1, 1]	1	W1
5	[1, -1, 1]	1	W2
5	[1, 1, -1]	1	W1
5	[1, 1, 1]	1	W0
6	[-1, -1, -1]	-1	W0
6	[-1, 1, -1]	-1	W2
6	[-1, -1, 1]	1	W1
6	[1, 1, -1]	1	W1
6	[1, 1, 1]	1	W0
6	[1, -1, -1]	-1	W3
6	[-1, 1, -1]	-1	W2
6	[-1, 1, 1]	1	W3

8- Summing the energy function for all 8 vectors containing a test sign is done with each weight sign vector stored in look up table.

$\sum$ Ep1= -24
$\sum$ Ep2= 4
According to the comparisons made between the test vectors of the sign with stored vectors in the look up table. **The first id (5)** stored in look up table have minim summing the energy function for all 8 vectors containing a test sign..

9 Calculate Confidence (Convergence Rate CR) Depending on The Energy Function using equation 3.1

CR1 = -24/24 *100= 100% ….. Identical
CR2= 4/24*100 = 16%

10 Returns alphabet corresponding to the input sign

5

Figure 4.1:case study steps

For each vector input, the final total of the energy functional is obtained, and the pattern with the lowest sum is selected from those that have been previously saved.

4.3 Real Time Hand Gesture Recognition RTHGR Evaluation

In this stage, the basic aspects of the stages of the proposed method for learning about many sign languages are taken separately. The assessment can be made on the basis of the first three points mentioned in the previous section

As mentioned previously, the data used in this research is American Sign Language, with an average of 126 images for each sign in the training stage and 54 images for each sign in the testing stage.

The dataset of alphabets from American Sign Language is separated into 26 volumes, representing the different categories. 87,000 images with a size of 200x200 pixels. training data set. 26 volumes are for letters A-Z. The test data is on 26 images only. Show (Figure 4.2). This dataset can be obtained by visiting this link: https://www.kaggle.com/datasets/grassknoted/asl-alphabet.

	
A	B

Figure 4.2: A Letter "A and B" From American Sign Language.

Both Chinese sign language and Arabic sign language were used in RTHGR system. With an average of 105 images for each sign in the training stage and 45 images for each sign in the testing stage. The data set for the numbers has been used with 52 images for each number in the learning stage and 23 images for each number in the testing stage. The proposed RTHGR uses a computer camera with a video quality of 720p (16:9, 30 fps).

4.3.1 Accuracy for Hand Gestures Recognition

The researchers have not previously discussed the use of associative memory in the real-time recognition of hand gestures. The reason is attributed to the need for associative memory to have a high storage capacity, so it has been used the MCA associative memory method solves a high storage capacity for recognizing hand gestures.

The confusion matrix is widely used in evaluating systems because accuracy, percentage, and other evaluation criteria depend on the results of the confusion matrix.

The accuracy criterion was used to evaluate the recognition of sign language for all the data sets used in the RTHGR system. Depending on the confusion matrix, determine the parameters of the accuracy criterion where:

A true negative (TN): is a sign that can be confidently classified as a negative sign.

True positive (TP): is the sign that actually belongs to the correct sign class.

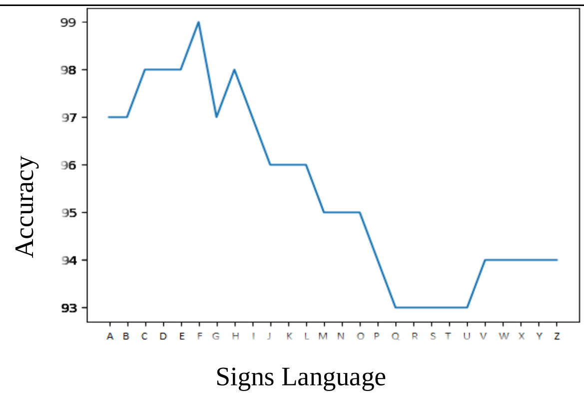
False positives (FP): are class-positive signs but are actually class-negative signs.

False negatives (FN): are negative signs but are actually positive signs.

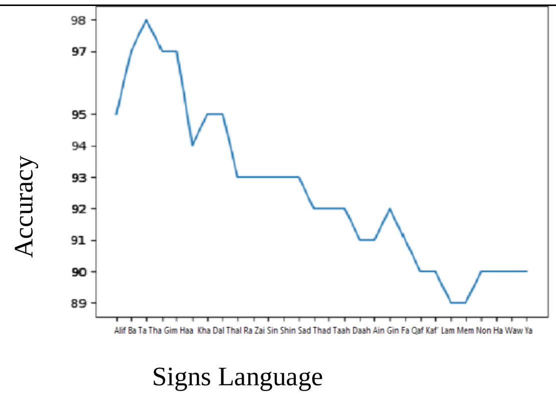
Equation 2.4 was used to extract the accuracy of the RTHGR system.

The proposed method gave good results in real-time with an accuracy of 95.42% when using ASL, It was also used on multiple datasets, each time giving promising results. Recognition of Chinese sign language achieved 92.13 % accuracy, while recognition of Arabic sign language achieved 93.55 %

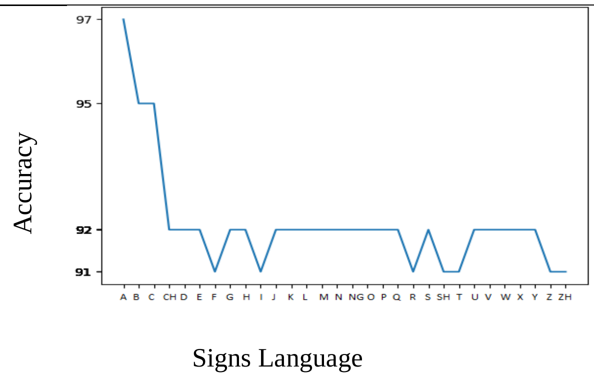
accuracy. Overall, the number recognition accuracy was 94 %. The accuracy result of the proposed system in real-time is illustrated in (Figure 4.3).



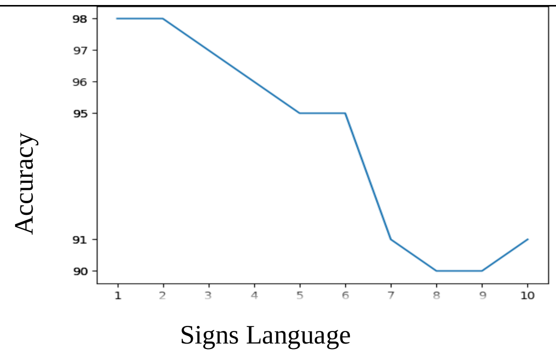
American Sign Language Accuracy = 95.42
26 Sign



Arabic Sign Language Accuracy = 93.55
28 Sign



Chinese Sign Language Accuracy = 92.13
30 Sign



Number 1-10 Accuracy = 94
10 Sign

Figure 4.3: Accuracy for Hand Gestures Recognition.

4.3.2 RTHGR Efficiency

The main advantage for RTHGR lies in taking advantage of the concept of orthogonality in MCA associative memory neural network and thus ensuring the use of the least number of weight matrices, which are only four. Therefore, during the retrieval process, the comparison process is easier. Using MCA reduced the traditional pre-processors, feature extraction, and solving the correlation problem between similar patterns.

The convergence phase in the RTHGR network gives impressive results in real-time, due to assignment only, and there are no complicated calculations or operations that take time as a primary factor. What distinguishes the proposed method for hand gesture recognition, the network is only one layer, the weights are restricted to just four, and the number of images used in the system is small compared to what CNN processes. The energy function was used to calculate the degree of confidence. Because the recognized pattern is not 100% identical with the saved sign patterns in the look-up table, a convergence rate of 70% to 90% or less shows that the convergence process is successful this is done without any complication due to the assignment process for only four weights per sign pattern. The convergence rate is shown in (Step 11) of the algorithm 3.3 hand gestures detector. The procedure of recognizing four signs with a confidence rating for both is depicted in (Figure 4.4).

One of the challenges facing sign language recognition systems is that even an expert person cannot always repeat the same sign with the same structure exactly, especially for complex signs with finger movement. Therefore, the possibility of deleting the wrong letter (sign) displayed on the translation screen gave the desired flexibility. The accuracy results of the proposed system within a real-time perspective will be discussed in the next section.

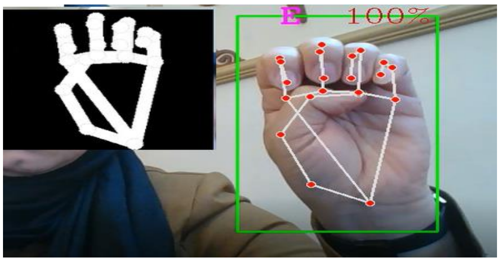
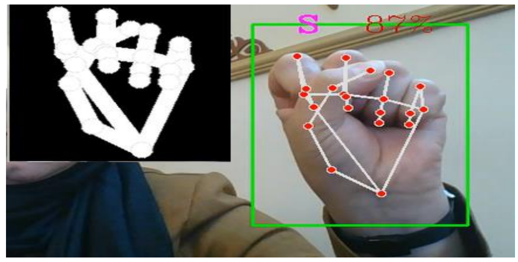
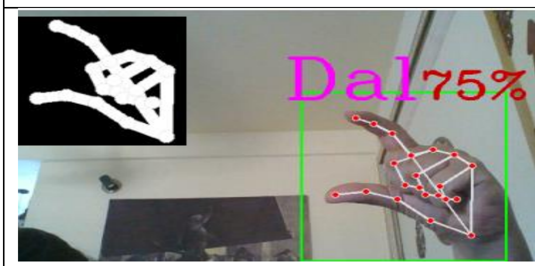
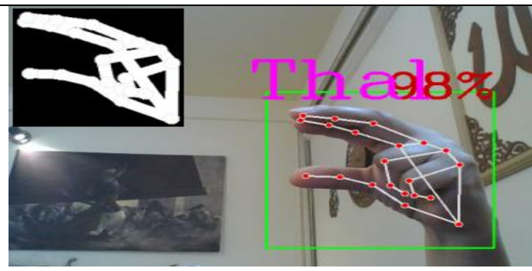
	
Recognize The Letter *E* with Convergence Rate 100%	**Recognize The Letter *S* with Convergence Rate 87%**
	
Recognize The Letter *Dal* with Convergence Rate 75%	**Recognize The Letter *Thal* with Convergence Rate 98%**

Figure 4.4: Recognize the Signs with Convergence Rate.

4.3.3 Evaluation of The Proposed System in Real-Time.

In the proposed RTHGR, the model has been trained on 126 images and tested on 54 images for each alphabetic letter. As a result, only 4,680 images were incorporated into the model regarding ASL. 70% of the data was allocated for training and 30% for testing. Both Chinese sign language and Arabic sign language were used in this proposed. With an average of 105 images for each sign in the learning stage and 45 images for each sign in the testing stage. The data set for the numbers has been used with 52 images for each number in the learning stage and 23 images for each number in the testing stage. Table 4.1 explains the result of the proposed method in real-time.

Table 4.1: The Result of The Proposed Method in Real-Time.

Data set	No. of train images per sign	No. of tested images per sign	No. of sign or class	The total number of test images used	Time of test images used	Accuracy
ASL	126	**54**	26	1040	0.36/s	95.42%
ArSL	105	45	28	1400	0.39/s	93.55%
ChSL	105	45	30	1500	0.40/s	92.13%
NSL	52	23	10	250	0.12/s	94%

The normal rate for video frames is 30 frames per second, so if the classes of signs are few, the number (period) of frames can be low. On the contrary, if the number of classes of signs is large, it needs time for bigger processors, so give a number (period) of frames that can be large. The possibility of choosing a frame at different speeds and within the actual time of the process of selecting the letter corresponding to a specific sign gives the translation process the necessary flexibility. It is worth noting that the translation process was applied to ASL, and the right and left hands were used.

The number of classes allocated for the right hand to recognize it is 26, while the number of classes that the left hand must recognize is only 10. The proposed system characterized is far from complicated, and the number of data sets used is much less than what researchers used in other ways, such as CNN. The proposed method of sign language recognition has been compared with some related research and the result of the comparison is presented in the table (4.2).

Table 4.2: Comparison of The Proposed Method with Other Methods

Reference	*Year*	*Approach*	*Accuracy*	*No of sign*	*Dataset*				*Data set*	*Real time*	*Execution time*
					ASL	ArSL	ChSL	No.			
[76]	**2018**	**Generalized Regression Neural Network**	**90.44%**	**26**	✓				**900**	**N/L**	**N/L**
[77]	**2018**	**Histograms of Oriented Gradients (HOG) and SVM**	**63.5%**	**30**		✓			**900**	**N/L**	**N/L**
[78]	**2018**	**Densely Connected Convolutional Neural Networks (DenseNet)**	**90.3**	**26**	✓				**64790**	✓	**0.12-0.14 ms**
[38]	**2018**	**MobileNet models**	**95.03**	**26**	✓				**87000**	✓	**2.42 s**
[40]	**2019**	**six-layer convolutional neural network**	**88.10 ± 1.48%**	**30**			✓		**12600**	**N/L**	**N/L**
[58]	**2020**	**Neutrosophic and fuzzy c-means**	**91.0%**	**28**		✓			**300**	**N/L**	**N/L**
[59]	**2020**	**Deep Convolutional Neural Network**	**94.34%**	**26**	✓				**12500**	✓	**N/L**
[60]	**2021**	**CNN And Long Short-Term Memory -LSTM**	**85.6%**	**28** **11**		✓		✓	**60000**	✓	**N/L**
[29]	**2021**	**Convolution Neural Network**	**90%**	**28**		✓			**N/L**	✓	**N/L**
[20]	**2022**	**Lightweight Convolutional Neural Network**	**94%**	**28**		✓			**5400**	✓	**N/L**
[61]	**2023**	**3D convolutional neural networks**	**91 ± 07**	**20**		✓			**3700 vido**	**N/L**	**N/L**
The Proposed Method	**2023**	**Landmark point and Multi-Connect Architecture Associative Memory**	**95.42%** **93.55%** **92.13%** **94%**	**26** **28** **30** **10**	✓	✓	✓	✓	**4680** **4200** **4500** **750**	✓	**0.12-0.40s**

4.4 Qualitative Evaluation

Qualitative evaluations of systems require gathering information through participant observation, in-depth interviews, comparison, and interpretation are all part of qualitative data analysis. A questionnaire was administered to many institutions offering services for the rehabilitation of the deaf and dumb to get this information. All details of the questionnaire will be discussed in this section.

4. 4.1 The Curriculum of The Study

An analytical descriptive approach to the proposed system as it exists in reality. It is concerned with an accurate description and is expressed in qualitative and quantitative terms, in order to investigate the various aspects of the application and its relationships with reality, will work on presenting and analyzing the results of the questionnaire, to reach conclusions.

4. 4.2 Study community

The questionnaire was based on targeting the group most related to and in contact with the deaf-dumb community. The questionnaire was conducted by sign language teachers (**%55**) and a social researcher (**%20**) as well as the administrative staff who deal with the deaf-dumb community (**%25**). Figure (4.5) show the study community according to the job title. The questionnaire was conducted on three institutes for the rehabilitation of the deaf and dumb in the governorates of Nineveh, Babylon and Najaf. and the number of the sample was 100.

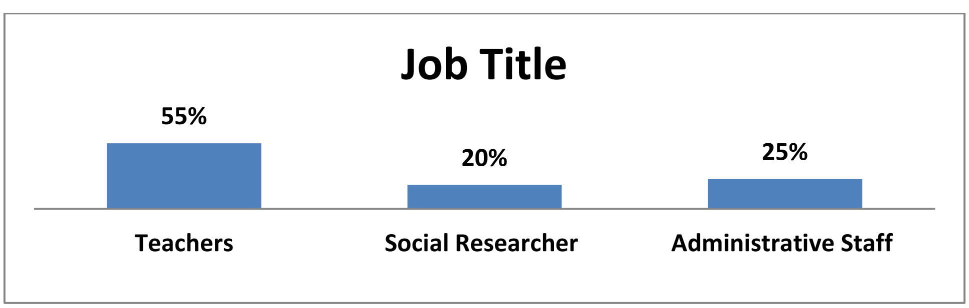


Figure 4.5: Population According to The Job Title.

4. 4.3 The Study Sample

The questionnaire has been consisted of (100) records and was taken from different people and different specializations and in varying proportions working within the field of dealing with sign language. 43% of the sample was for the Directorate of Social Welfare / Department of Care for People with Special Needs in Nineveh, as well as 40% of the sample was for the Al-Amal Institute for the Deaf and Dumb in Najaf, and finally, 17% of the sample was for the Al-Amal Institute for the Deaf and Dumb in Babylon, (Figure 4.6) show distribution of sample by place of work. It is worth noting that the aforementioned departments are with the Iraqi Ministry of Labor and Social Affairs.

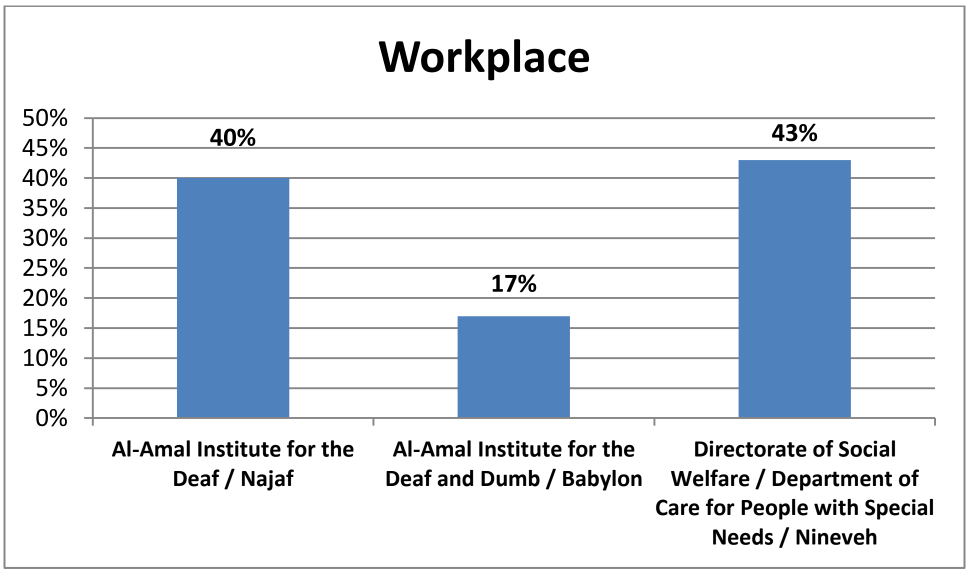


Figure 4.6: Distribution of Sample by Place of Work.

4. 4.4 Study Tool

The questionnaire was used as a study tool, specifically designed to determine the importance of a real-time sign language recognition system for the deaf and the dumb community.

4.4.5 Questionnaire Components

All considerations to the performance of the proposed system must be addressed by the questionnaire. Hence, to show that the research has achieved its main goal. Accordingly, the questionnaire consists of 7 pivots that include:

1. **The First Pivot: personal information and contains six important information.**

   1- Age: 25-40 years, 40-60 years

   2- Gender: Male, Female

   3- Job:

   4- Functional service:

   5- Do you have knowledge of sign language?

   6- Do you practice sign language in communicating with the deaf and dumb community?

2. **The Second Pivot: the pivot of the interactive interface of the system and contains four questions.**

   1- The main interface is easy to use?

   2- Is the required information (sign) enough to be entered into the program interface?

   3- The data that is displayed in the main interface is useful and sufficient to communicate the desire or need of the person who is deaf or dumb?

   4- Do you think that providing the opportunity to implement sign language interpretation and translation programs is a correct choice?

3. **The Third Pivot: the pivot of the proposed system for learning about sign language, and it contains five questions.**

   1- Is the proposed system compatible with the requirements of people who use sign language?

   2- Is the proposed system easy to learn?

3- The proposed system is characterized by the possibility of using it by persons who use sign language to communicate with their peers.

4- The proposed system is characterized by the necessary flexibility, especially with the possibility of erasing the wrong letter?

5- Do you support the importance of such projects to alleviate communication problems between the deaf and dumb community and ordinary people?

4. **The Fourth Pivot: the pivot of the accuracy of the system and contains five questions.**

1- Is the information generated from the system characterized by the accuracy and flexibility necessary to correct the error, if any?

2- Do you have confidence in the results obtained through the proposed system in the interpretation and translation of sign language?

3- Does the proposed system contribute to providing alternative options for communication?

4- What do you think about the presented results of the proposed system in interpreting the sign in real time?

5- Do you think that the results displayed in real time are of real benefit?

5. **The Fifth Pivot: the pivot of decision-making speed and contains four questions.**

1. The proposed system helps in interpreting the sign between the deaf and dumb community and ordinary people?

2. The proposed system provides alternatives and solutions quickly for communication?

3. The deaf and dumb community face difficulty in communicating with others without relying on systems such as the proposed system?

4. The deaf and dumb community needs a sign translation system that meets the urgent needs of communicating with others?

6. **The Sixth Pivot: The system cost pivot is one of the hardware requirements and contains three questions.**
   1. The proposed system does not need additional capabilities and is satisfied with a mobile camera or a computer camera?
   2. Do you think that the acquisition of devices that implement the system requires exorbitant costs?
   3. Is there a computer available in the deaf and dumb centers that can be used to learn signs while giving lessons to the deaf and dumb community?
7. **The Seventh Pivot: the pivot of applicability, which contains four questions.**
   1- Do you agree that the idea of the program is applicable?
   2- Do you think that the idea of research eliminates the need to use writing with paper and pen and the possibility of translating the needs of the deaf and dumb by using electronic technologies such as mobile devices?
   3- Do you support the idea that the results of such programs can help people with disabilities who are deaf and dumb to be involved in regular governmental educational centers (schools and universities)?
   4- Do you agree that the outputs of the proposed system help the deaf and dumb community to manage their affairs in state departments or courts during trials?

4.4.6 Staging Scale

Study participants' ratings on a scale from one to five were considered valid. The analysis used a five-point Likert scale, with one being the lowest score and five the highest. Table (4.3) displays the grading scale that was used.

Table (4.3): Relative Weight Scale.

Relative weight	***1***	***2***	***3***	***4***	***5***
Qualitative	Very low	Low	Medium	High	Very High

4.4 Results of The Field Study

In addition to what dealt with in the previous section of analyzing and showing the percentages of the places where the questionnaire was conducted, as well as the job titles of the sample community. will show in this section it includes distributing the sample answers according to some personal variables, answering the study questions, and presenting the results.

4.5.1 Statistical Description of The Study Sample

The following sections show a description of the study sample according to the variables: gender, qualification to know sign language and years of service.

4.5.2 Study Sample According to The Gender Variable

Figure 4.7 shows the distribution of the study sample according to the gender variable.

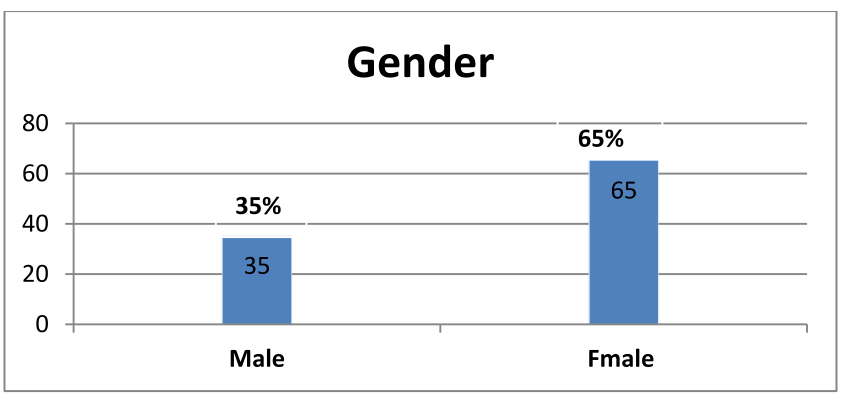


Figure 4.7: Study Sample According to The Gender Variable.

Figure 4.7 shows that the percentage of females outweighs the males. explain the discrepancy between males and females to the percentage of females employed as sign language teachers who are originally much more than males. The reason is that the questionnaire focused on the most category that deals with sign language to connect with the deaf community, and they are teachers, and most of the teachers are females, so see that the percentage of females is greater.

4.5.3 The study sample according to the knowledge of sign language

The questionnaire shows the qualification to know sign language (Figure 4.8) shows the distribution of the study sample according to the qualification variable:

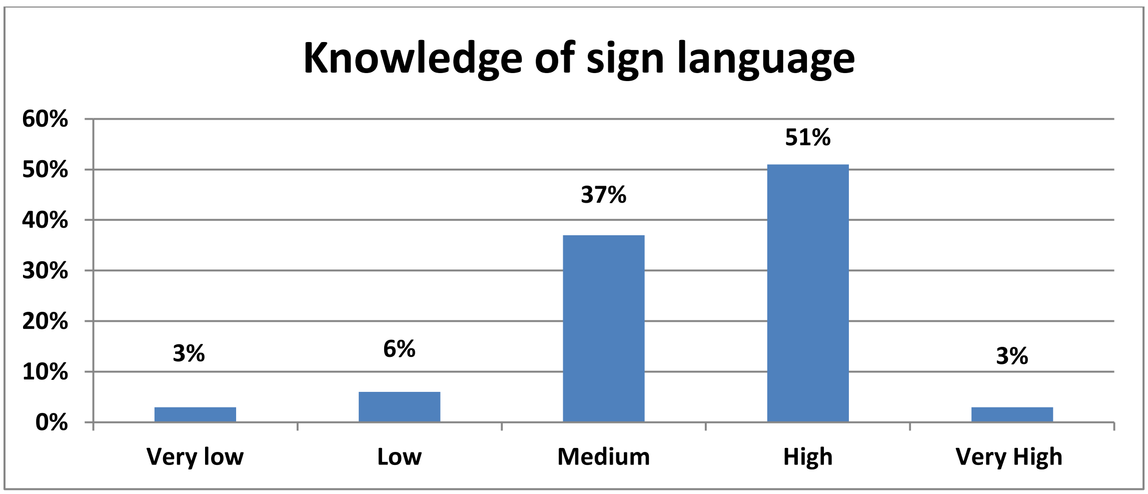


Figure 4.8: Study Sample According to Knowledge of Sign Language.

As shown in (Figure 4.8) the questionnaire was taken from various experts and different specialties with varying rates, and that the percentage. It is worth mentioning that all those in the centers where the questionnaire was conducted are familiar with sign language because of their direct contact with children who are deaf or dumb. As for the difference in knowledge, it indicates the long experience in practicing sign language. Or their practice of teaching sign language for a longer period than others led to the accumulation of experiences for them.

4.5.4 A study sample according to the job title

According to (Figure 4.5). % 55 of the study sample, the job title is teacher, % 20 social researcher", and 25% are administrative employees in institutes for the deaf and dumb,

Note that all the study samples have a close relationship with the deaf-dumb community, and they are the ones who give their opinion on the application explicitly and honestly because of their direct contact with the deaf-dumb community.

4.5.5 Study sample according to years of service variable

Figure (4.9) Distribution of the study sample according to the years of service variable:

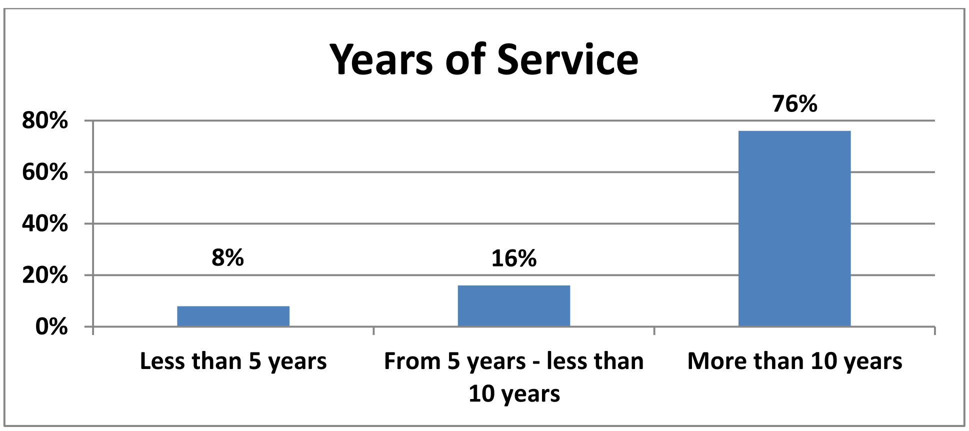


Figure 4.9: Sample According to Years of Service Variable.

In (Figure 4.9), 76% of the study sample individuals have reached their years of service "10 years or more" which is a high percentage, 8% "less than 5 years" and 16% from 5 years to less than 10 years. In addition to scientific qualification in specialization to know sign language, years of service add more experience in knowing sign language; This is due to the fact that those who occupy specialized positions must have sufficient years of experience, especially for higher management categories, teacher, and the requirements for occupying positions in government institutions that deal with the deaf and dumb community.

They must have experience teaching scientific subjects through signs. It is worth mentioning that all primary school curricula are taught to deaf children using signs. , which is a clear indication that those who have served for more than 10 years have A high percentage and they have the skill, scientific experience, and specialization, which confirms the sobriety and importance of the questionnaire.

4.5.6 Statistical Description of the RTHGR System

According to the 100 questionnaires which represent the study sample, the following sections will be analyzed and discussed with all pivots respectively. The Average, standard error, relative weight, and descriptive degree has been calculated. Average represents a value around which other values gather and the extent of their balance. This means the standard error of the questionnaire results should be small.

4.5.6.1 Second Pivot: Interactive Interface of The System.

In general, the (Figure 4.10) shows the sample's answers to the questionnaire for the second pivot, which included four questions related to the interactive interface of the system.

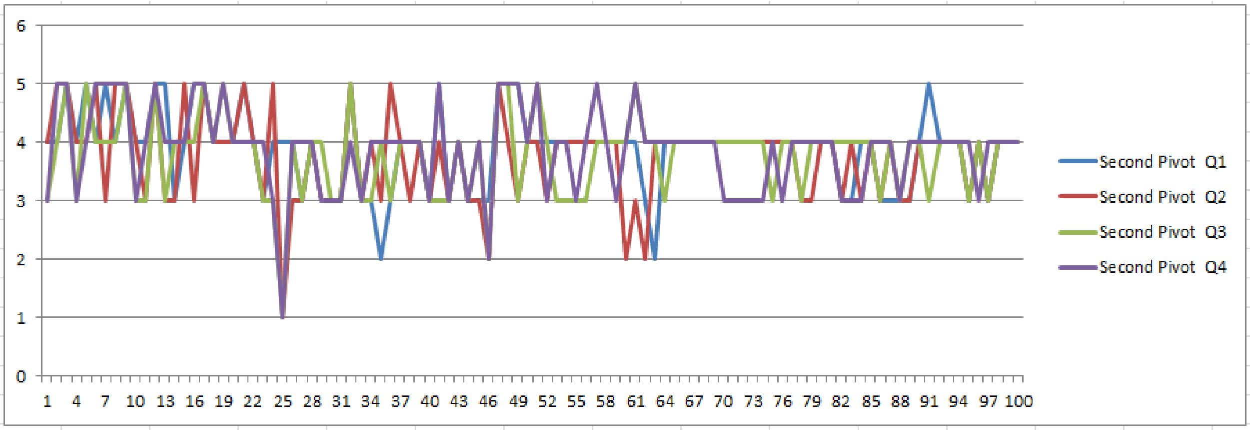


Figure 4.10: Rate of Answers on The Second Pivot.

Table (4.4) shows the Average, standard error, relative weight, and descriptive degree of the system interface pivot.

Table 4.4: Analysis of Pivot of Interactive Interface.

NO.	The Paragraphs	Standard Error	Average	Relative weight	Qualitative
1	**The main interface is easy to use?**	*0.2*	*3.9*	*4*	*High*
2	**Is the required information (sign) enough to be entered into the program interface?**	*0.2*	*3.7*	*4*	*High*
3	**The data that is displayed in the main interface is useful and sufficient to communicate the desire or need of the person who is deaf or dumb?**	*0.2*	*3.8*	*4*	*High*
4	**Do you think that providing the opportunity to implement sign language interpretation and translation programs is a correct choice?**	*0.2*	*3.9*	*4*	*High*

It is clear from Table (4.4) that question No. (1), (2), (3), and (4) their standard error is low (equal to 0.2) the relative weight is 4 which comes to a "high" weight degree.

Accordingly, questions number (1), (2), and (3) emphasized that the proposed system is easy to use and flexible interface for dealing with sign language, as it facilitates the process and shows the results faster and the required information (sign) is enough to be entered into the program interface. And the question (4) confirms that choosing a program that recognizes sign language is one of the suitable options.

According to the second pivot of all questions, the average was close to (4), indicating that the interactive interface for the proposed system is easy and the standard error according to results is (0.2), which added more credibility and reliability to the pivot's questions.

4.5.6.2 The Third Pivot: The Proposed System for Learning About Sign Language

In general, the (Figure 4.11) below shows the sample's answers to the questionnaire for the third pivot, which included five questions related to the proposed system to learn about sign language.

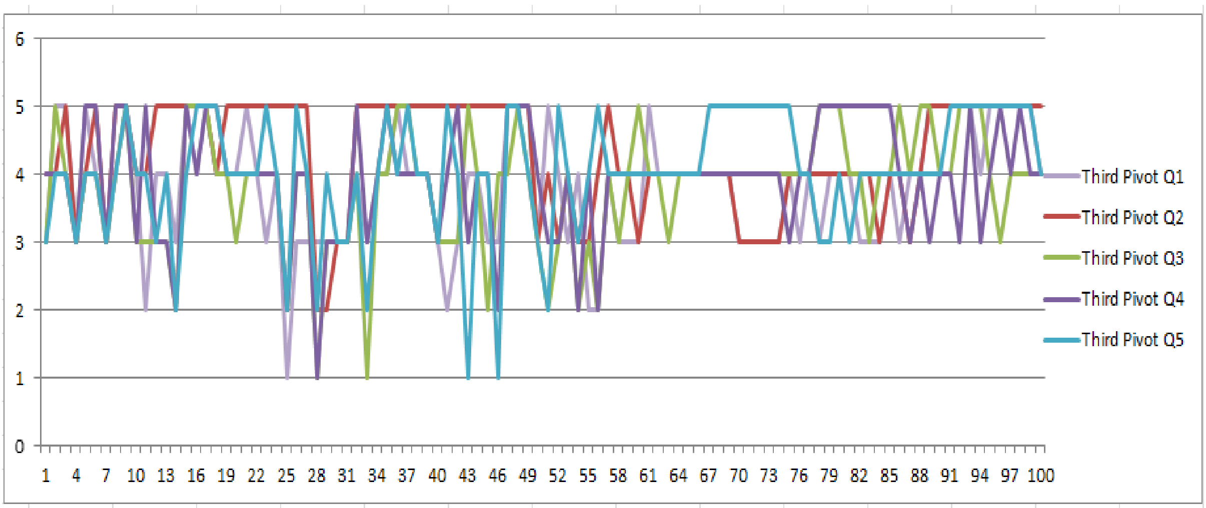


Figure 4.11: Rate of Answers on The Third Pivot.

Table (4.5) shows the Average, standard error, relative weight, and descriptive degree of the proposed system for learning about sign language pivot

Table 4.5 Analysis the Pivot of the proposed system for learning about sign language

NO.	The Paragraphs	Standard Error	Average	Relative weight	Qualitative
1	**Is the proposed system compatible with the requirements of people who use sign language?**	*0.2*	*3.8*	*4*	*High*

2	**Is the proposed system easy to learn?**	***0.2***	***4.3***	***5***	***Very High***
3	**The proposed system is characterized by the possibility of using it by persons who use sign language to communicate with their peers**	***0.2***	***3.8***	***4***	***High***
4	**The proposed system is characterized by the necessary flexibility, especially with the possibility of erasing the wrong letter?**	***0.2***	***3.9***	***4***	***High***
5	**Do you support the importance of such projects to alleviate communication problems between the deaf and dumb community and ordinary people?**	**0.2**	***4***	***4***	***High***

For all questions, the standard error was (0.2) and a high rate for average. Question No. (1) confirmed that the proposed system is commensurate with the requirements of people who use sign language. Likewise, questions (2) and (4) gave a very high percentage, emphasizing the ease of learning and the flexibility that characterizes the system, especially with the possibility of modifying the sentence that will be written by adding the feature of erasing the wrong letter that may appear during writing the translation for the meaning of the sign. The (3) and (5) questions emphasized the importance of application by those who use sign language with a high percentage, as well as the possibility of alleviating the communication problems they face with others. From the above table (4.5) and Figure (4.11), conclude that a ***high*** percentage of the questionnaire sample emphasized giving the third pivot, which is related to the sign language recognition system, its importance for the deaf and dumb community, and the possibility of alleviating communication problems with others by applying programs that translate the sign into sentences.

4.5.6.3 The Fourth Pivot: The Accuracy of The System.

In general, the (Figure 4.12) below shows the sample's answers to the questionnaire for the fourth pivot, which included five questions related to the degree of the accuracy of the system pivot.

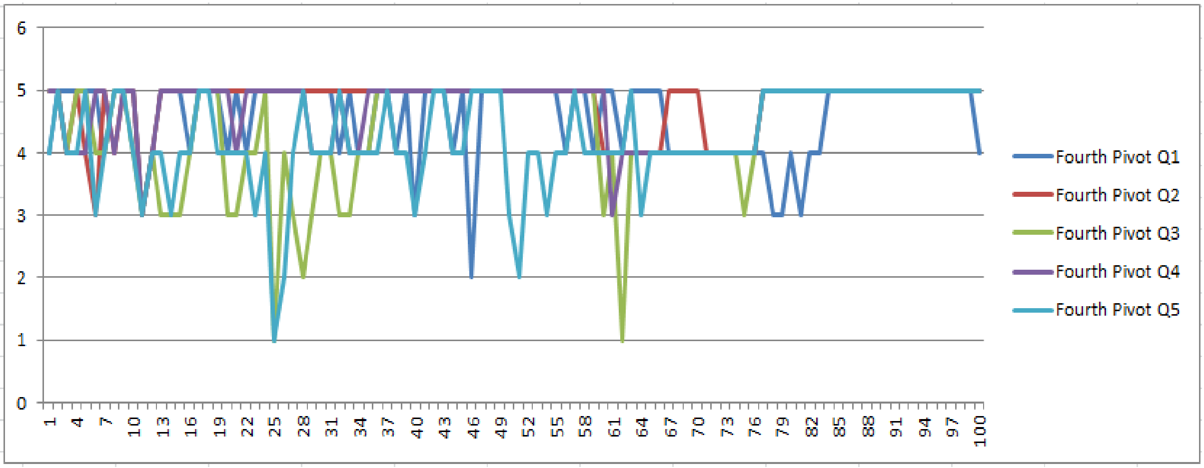


Figure 4.12: Rate of Answers on The Fourth Pivot.

Table (4.6) shows the Average, standard error, relative weight, and descriptive degree of the fourth pivot of the accuracy of the system.

Table 4.6 Analysis of Pivot of the accuracy of the system**.**

NO.	The Paragraphs	Average	Standard Error	Relative weight	Qualitative
1	**Is the information generated from the system characterized by the accuracy and flexibility necessary to correct the error, if any?**	***4.6***	***0.1***	***5***	***Very high***
2	**Do you have confidence in the results obtained through the proposed**	***4.8***	***0.1***	***5***	***Very high***

	system in the interpretation and translation of sign language?				
3	**Does the proposed system contribute to providing alternative options for communication?**	***4.4***	***0.2***	***5***	***Very high***
4	**What do you think about the presented results of the proposed system in interpreting the sign in real time?**	***4.7***	***0.1***	***5***	***Very high***
5	**Do you think that the results displayed in real time are of real benefit?**	***4.3***	***0.2***	***4***	***high***

The fourth pivot, which is related to the accuracy of the proposed system, as the questionnaire sample gave very high rates, a standard error is (0.1 - 0.2), and a very high relative weight. Where question (1) confirmed that the information generated from the system is characterized by the accuracy and flexibility necessary to correct the error, if any. As well as the (2) and (3) questions, they emphasized the reliability of the results obtained through the proposed system in the interpretation and translation of sign language, and that the proposed system contributes to providing alternative options for communication. Both the (4) and (5) questions gave their results a high relative weight relative to the importance of the time component in the presented results of the proposed system in interpreting the sign in real-time, and the results presented in real-time are of real benefit.

According to the fourth pivot of all questions, the average was close to (5), indicating that the accuracy of the proposed system is very high and the standard error according to results is (0.1-0.2), which added more credibility and reliability to the pivot’s questions.

4.5.6.4 The Fifth Pivot: Decision-Making Speed.

In general, the (Figure 4.13) below shows the sample's answers to the questionnaire for the fifth pivot, which included four questions related to the decision-making speed of the proposed system.

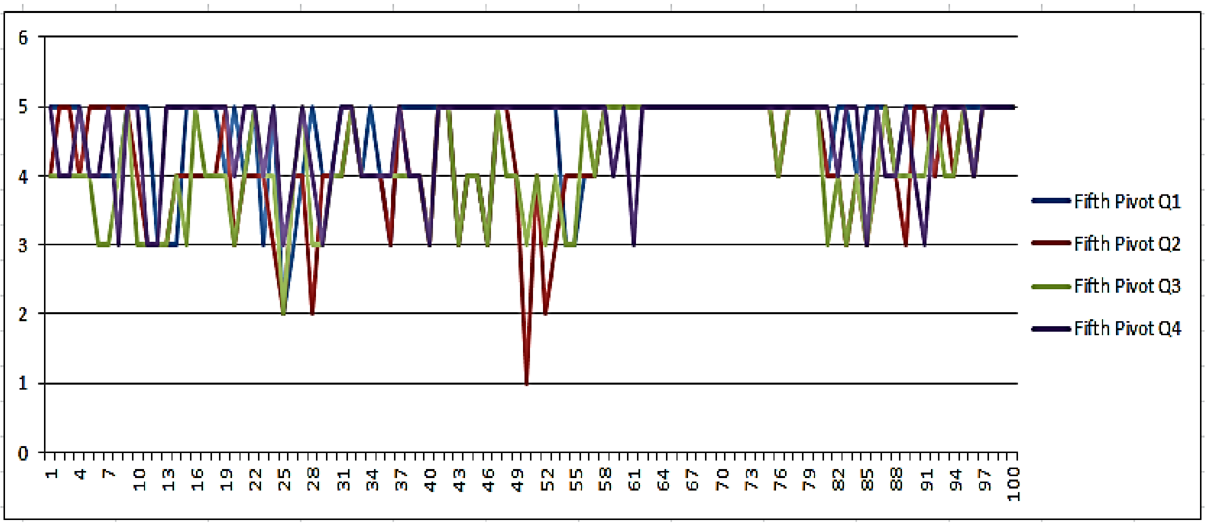


Figure 4.13: Rate of Answers on the fifth Pivot.

Table (4.7) shows the Average, standard error, relative weight, and descriptive degree of pivot: of decision-making speed.

Table 4.7: Analysis of Pivot of decision-making speed **.**

NO.	The Paragraphs	Standard Error	Average	Relative weight	Qualitative
1	**The proposed system helps in interpreting the sign between the deaf and dumb community and ordinary people.**	*0.2*	*4.6*	*5*	*Very high*
2	**The proposed system provides alternatives and solutions quickly for communication.**	*0.2*	*4.5*	*5*	*Very high*

3	**The deaf and dumb community faces difficulty in communicating with others without relying on systems such as the proposed system?**	***0.2***	***4.2***	***4***	***High***
4	**The deaf and dumb community needs a sign translation system that meets the urgent needs of communicating with others?**	***0.1***	***4.6***	***5***	***Very high***

The fifth pivot, related to the accuracy of the proposed system, as the questionnaire sample gave very high rates, a standard error between (0.1 - 0.2), and a very high relative weight. Where question (1) confirmed that the system helps to interpret the sign between the deaf-dumb community and ordinary people, as is the case with regard to question (2), where the questionnaire sample confirmed that the proposed system provides alternatives and speed communication between the deaf-dumb community and original people, and based on the (1) and (2) questions, the result of the two questions was confirmed The (4) and (5) is that the deaf and dumb community greatly needs systems that translate sign to meet their desires while communicating with others.

According to the fifth pivot of all questions, the average was close to (5), indicating that the decision-making speed of the proposed system is very high and the standard error according to results is (0.1-0.2), which added more credibility and reliability to the pivot’s questions.

4.4.6.5 The Sixth Pivot: Cost of The Hardware Requirements

In general, the (Figure 4.14) below shows the sample's answers to the questionnaire for the Sixth pivot, which included three questions related to cost of the hardware requirements.

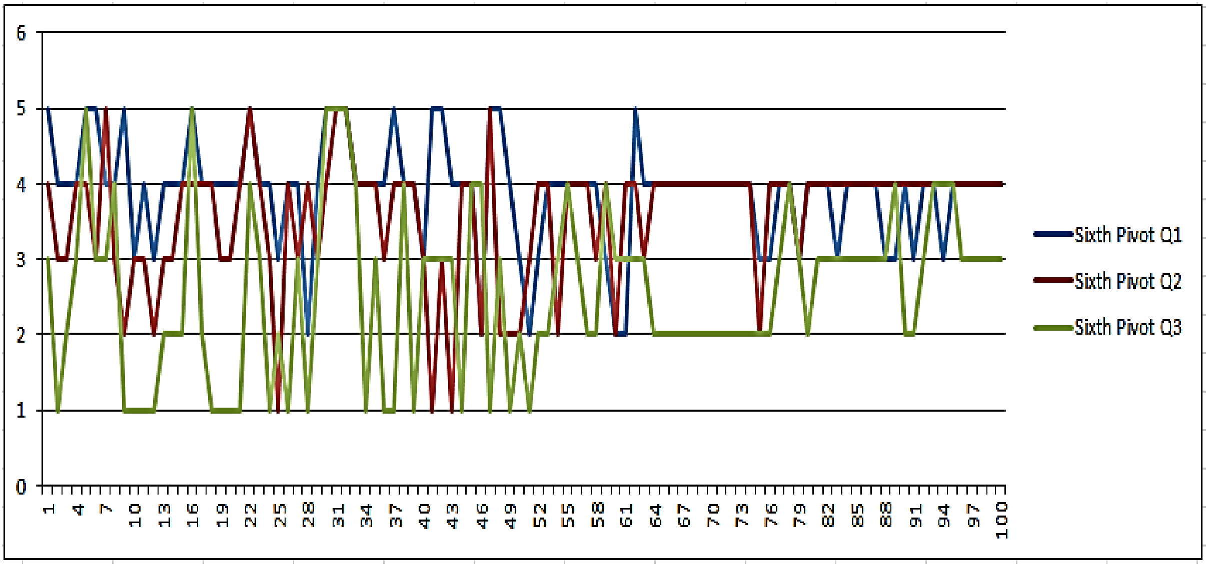


Figure 4.14: Rate of Answers on The Sixth Pivot.

Table (4.8) shows the Average, standard error, relative weight, and descriptive degree of pivot: cost of the hardware requirements.

Table 4.8: Analysis cost of the hardware requirements pivot**.**

NO.	The Paragraphs	Average	Standard Error	Relative weight	Qualitative
1	**The proposed system does not need additional capabilities and is satisfied with a mobile camera or a calculator camera?**	***3.9***	***0.2***	***4***	***High***
2	**Do you think that the acquisition of devices that implement the system requires exorbitant costs?**	***3.6***	***0.2***	***4***	***High***
3	**Is there a computer available**	***2.6***	***0.4***	***3***	***Medium***

	in the deaf and dumb centers that can be used to learn signs while giving lessons to the deaf and dumb community?				

In the sixth pivot, there was a difference of opinion, as the relative weight of the pivot as a whole ranged between (3-4) and that was based on the infrastructure of the institution and its computer equipment. Where the standard error for the first question was (0.2), which states that the proposed system does not need additional hardware and suffices with a mobile or a laptop, and the relative weight of this question was (3)

As for the (2) and (3) questions, there was a clear discrepancy in opinions, especially since most of the centers do not have computers in abundance, the infrastructure of the three centers had an impact on the great discrepancy in opinions, especially since the center for the deaf and dumb in Nineveh province was bombed during the war on ISIS, and it gave results questionnaire ranks high because of the need to acquire computer hardware that implements the system. As for the availability of a computer in deaf and dumb centers that can be used to learn signs while giving lessons to the deaf and dumb community, the answers also vary according to the infrastructure of each center separately.

4.5.6.6 The Seventh Pivot: of Applicability

In general, the (Figure 4.15) below shows the sample's answers to the questionnaire for the seventh pivot, which included four questions related to the applicability of the proposed system.

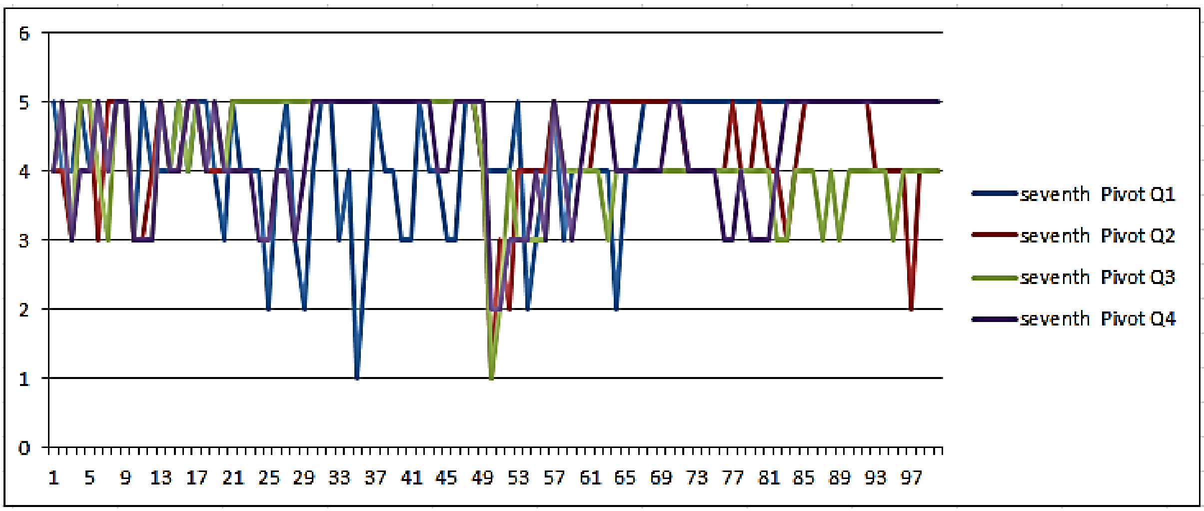


Figure 4.15: Rate of Answers on The Seven Pivot.

Table (4.9) shows the Average, standard error, relative weight, and descriptive degree of pivot: of applicability

Table 4.9 Analysis of Pivot of applicability**.**

NO.	The Paragraphs	Average	Standard Error	Relative weight	Qualitative
1	**Do you agree that the idea of the program is applicable?**	***4.3***	***0.2***	***5***	***Very high***
2	**Do you think that the idea of research eliminates the need to use writing with paper and pen and the possibility of translating the needs of the deaf and dumb by using electronic technologies such as mobile devices?**	***4.4***	***0.2***	***5***	***Very high***
3	**Do you support the idea that the results of such programs can help people with disabilities who are deaf and dumb to be involved in regular governmental educational centers (schools and universities)?**	***4.2***	***0.2***	***4***	***High***

4	**Do you agree that the outputs of the proposed system help the deaf and dumb community to manage their affairs in state departments or courts during trials?**	***4.3***	***0.2***	***5***	***Very high***

From the above table (4.9) the questionnaire gave a very high value to the seventh pivot, which is related to the very high applicability, as the sample confirmed very highly that the idea of the proposed system is applicable, and the standard error for the question (1) is (0.2). According to the (2), (3), and (4) questions, which gave a standard error (0.2), infer that the proposed system is very highly applicable, which will help translate the needs of the deaf-dumb, as the results of such the proposed system can help deaf-dumb people with disabilities to engage in regular government educational centers (schools and universities) and help the deaf and dumb community to manage their affairs in many places or courts during trials.

Finally, based on the tables and the sample that conducted the questionnaire, as well as the Figures (4.5), (4.8) and (4.9) that confirm the sample, a large percentage of people had long experience in rehabilitation centers for the deaf and dumb, and they have very high reliability and credibility to evaluate the proposed system within the seven pivots. The results questionnaire gave very satisfactory results, with a standard error between (0.1 - 0.2) for all pivots.

4.6 Summary

In this chapter, the results reached by the proposed system are discussed and taken into consideration the discussion of the results is according to the contributions made by the proposed system. The success of the proposed method RTHGR for the sign language recognition system is attributed, to the use of the landmark point in the detecting and segmentation phase, giving it structure to the shape of the hand and in different environments, as well as the use of associative memory as a new direction by MCA with more efficiency and small network size. Three neurons and their weights are constant, which are four based on the concept of Orthogonality, all these features made RTHGR able to avoid the limitations in the convergence phase. A questionnaire was also conducted to evaluate the system as a whole and the last stage in particular, and the method gave satisfactory and promising results.

CHAPTER FIVE

CONCLUSIONS AND FUTURE WORK

5.1 Conclusion

In recent years, hand gestures' popularity has increased to use in many areas, and they have begun to appear in a wide range of useful and even crucial contexts. Gestures are a form of non-verbal communication that helps deaf people to communicate with others. Therefore, there is an urgent need to develop applications capable of interpreting the meaning of the gesture by building an accurate and effective system. Gesture recognition within a computer vision approach is therefore an appropriate option. The proposed method gave results in real-time with 94%, 95%.42, 92.13%, and 93.55% accuracies for numbers and American, Chinese, and Arabic sign language sequentially. The conclusions of the proposed system can be limited to the following points:

1. Different images as frames of video were put to the test using a collection of ASL images and other data sets captured in different lighting environments. Using the landmark point is very efficient against changing the illumination in addition to avoiding the clutter in the background.
2. The environment in which the system operates is not specified, thus the proposed system will be able to adapt to any for any target environment in terms of lighting and background.
3. Using associative memory as a substitute for other methods in hand gesture recognition is a new direction, especially since associative memory does not work in real-time, so the use of MCA was a good choice because it reduced the weights that are used in the training and convergence processes to half within the concept of orthogonality.

4. The number of sign language patterns that were used in the training phase was less than other methods.
5. Using MCA provided the ability to deal with the correlation between similar hand gesture signs with high efficiency, which consequently led to the RTHGR system obtaining a high accuracy that reached above 95% for the different data sets.
6. The proposed system, through the use of MCA, gives the possibility of learning the new patterns of sign without affecting the previous learning process, will be an addition to knowledge only.
7. The ability to convert the signs into words or sentences is an important contribution that gives the proposed system its efficiency and ability to be applied.

5.2 Future Works

The findings of the proposed method using associative memory as a new direction with real-time sign language recognition showed promising results. Among the questionnaire that was conducted, we can summarize the future work according to the needs of the deaf and mute category to benefit from programs that contribute to conveying what they desire to ordinary people:

1- Studying development the proposed method to learn about the local sign language, as the Iraqi Ministry of Labor and Social Affairs has a guide for signs who can use and apply it, and thus benefit the deaf student.

2 - Studying to make the output from the system as a picture of the meaning of the sign (for example, entering the sign of the house, the output will be a picture of the house) or make the output from the system audio, not just text.

3 - Studying to make the communication from two sides possible to convey the desire of the average person as text or voice and translate it into signs understood by the deaf and dumb community and vice versa.

4 – Studying the combining the proposed system method with other artificial intelligence methods for increasing the accuracy of gesture recognition.

5- Studying the possibility of translating the sign language of one country into the sign language of another country.

REFERENCES

1. ADEYANJU, I. A.; BELLO, O. O.; ADEGBOYE, M. A. Machine learning methods for sign language recognition: A critical review and analysis. Intelligent Systems with Applications, 2021, 12: 200056
2. M. Yasen and S. Jusoh, “A systematic review on hand gesture recognition techniques, challenges and applications,” PeerJ Comput. Sci., vol. 5, p. e218, 2019.
3. CAMGOZ, Necati Cihan, et al. Sign language transformers: Joint end-to-end sign language recognition and translation. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2020. p. 10023-10033.
4. TIAN, Ling, et al. Video big data in smart city: Background construction and optimization for surveillance video processing. Future Generation Computer Systems, 2018, 86: 1371-1382.
5. SARMA, Debajit; BHUYAN, M. K. Methods, Databases and Recent Advancement of Vision-Based Hand Gesture Recognition for HCI Systems: A Review. SN Computer Science, 2021, 2.6: 1-40
6. Perimal M., Basah S.N., Safar M.J.A., Yazid H. Hand-Gesture Recognition-Algorithm based on Finger Counting. J. Telecommun. Electron. Comput. Eng. 2018;10:19–24
7. SALUNKE, Tejashree P.; BHARKAD, S. D. Power point control using hand gesture recognition based on HOG feature extraction and K-NN classification. In: 2019 International Conference on Computing Methodologies and Communication (ICCMC). IEEE, 2018. p. 1151-1155.
8. Himadri NS, Shinjini R, Sudipta S, Sayan T, Suhrid KC.. A machine learning based approach for hand gesture recognition using distinctive feature extraction. In: 2018
9. ZHU, Yuhui; JIANG, Shuo; SHULL, Peter B. Wrist-worn hand gesture recognition based on barometric pressure sensing. In: 2018 IEEE 15th International Conference on Wearable and Implantable Body Sensor Networks (BSN). IEEE, 2018. p. 181-184
10. TIAN, Zengshan, et al. WiCatch: A Wi-Fi based hand gesture recognition system. IEEE Access, 2018, 6: 16911-16923.

11. SAPIENZA, Stefano, et al. On-line event-driven hand gesture recognition based on surface electromyographic signals. In: 2018 IEEE international symposium on circuits and systems (ISCAS). IEEE, 2018. p. 1-5.
12. GUO, Yisheng, et al. Development and Application of Gesture Recognition System for Intelligent Robot. In: IOP Conference Series: Materials Science and Engineering. IOP Publishing, 2018. p. 042172.
13. OLIVEIRA, Marlon, et al. Hand orientation redundancy filter applied to hand-shapes dataset. In: Proceedings of the 2nd International Conference on Applications of Intelligent Systems. 2019. p. 1-5.
14. DE SMEDT, Quentin; WANNOUS, Hazem; VANDEBORRE, Jean-Philippe. Heterogeneous hand gesture recognition using 3D dynamic skeletal data. Computer Vision and Image Understanding, 2019, 181: 60-72.
15. Prakash J., Gautam U.K. Hand Gesture Recognition. Int. J. Recent Technol. Eng. 2019;7:54–59.
16. Ma X., Peng J. Kinect sensor-based long-distance hand gesture recognition and fingertip detection with depth information. J. Sens. 2018;2018:5809769. doi: 10.1155/2018/5809769.
17. Tsoli A., Argyros A.A. Joint 3D tracking of a deformable object in interaction with a hand; Proceedings of the European Conference on Computer Vision (ECCV); Munich, Germany. 8–14 September 2018; pp. 484–500.
18. Xi C., Chen J., Zhao C., Pei Q., Liu L. Real-time Hand Tracking Using Kinect; Proceedings of the 2nd International Conference on Digital Signal Processing; Tokyo, Japan. 25–27 February 2018; pp. 37–42.
19. HERNÁNDEZ, Irene. Automatic Irish sign language recognition. University of Dublin, Trinity College, 2018.
20. AlKhuraym, Batool & Ben Ismail, Mohamed Maher & Bchir, Ouiem. (2022). Arabic Sign Language Recognition using Lightweight CNN-based Architecture. International Journal of Advanced Computer Science and Applications. 13. 10.14569/IJACSA.2022.0130438.
21. AL-HAMMADI, Muneer, et al. Hand gesture recognition for sign language using 3DCNN. IEEE Access, 2020, 8: 79491-79509.
22. WU, Xiao Yan. A hand gesture recognition algorithm based on DC-CNN. Multimedia Tools and Applications, 2020, 79.13: 9193-9205

23. MAHMOOD, Mayyadah Ramiz; ABDULAZEEZ, Adnan Mohsin; ORMAN, Zeynep. Dynamic hand gesture recognition system for kurdish sign language using two lines of features. In: 2018 International Conference on Advanced Science and Engineering (ICOASE). IEEE, 2018. p. 42-47.
24. Malik J., Elhayek A., Stricker D. Structure-aware 3D hand pose regression from a single depth image; Proceedings of the International Conference on Virtual Reality and Augmented Reality; London, UK. 22–23 October 2018; pp. 3–17.
25. Ge L., Ren Z., Li Y., Xue Z., Wang Y., Cai J., Yuan J. 3d hand shape and pose estimation from a single rgb image; Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; Long Beach, CA, USA. 15–20 June 2019; pp. 10833–10842
26. Devineau G., Moutarde F., Xi W., Yang J. Deep learning for hand gesture recognition on skeletal data; Proceedings of the 2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018); Xi'an, China. 15–19 May 2018; pp. 106–113
27. Chung H., Chung Y., Tsai W. An efficient hand gesture recognition system based on deep CNN; Proceedings of the 2019 IEEE International Conference on Industrial Technology (ICIT); Melbourne, Australia. 13–15 February 2019; pp. 853–858.
28. Bao P., Maqueda A.I., del-Blanco C.R., García N. Tiny hand gesture recognition without localization via a deep convolutional network. IEEE Trans. Consum. Electron. 2019 ;63:251–257. doi: 10.1109/TCE.2019.014971.
29. Rwelli, Rady & Shahin, Osama & Taloba, Ahmed. (2021). Gesture based Arabic Sign Language Recognition for Impaired People based on Convolution Neural Network. International Journal of Advanced Computer Science and Applications. 12. 10.14569/IJACSA.2021.0121273.
30. Nguyen X.S., Brun L., Lézoray O., Bougleux S. A neural network based on SPD manifold learning for skeleton-based hand gesture recognition; Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; Long Beach, CA, USA. 15–20 June 2019; pp. 12036–12045
31. LAI, Kenneth; YANUSHKEVICH, Svetlana N. CNN+ RNN depth and skeleton based dynamic hand gesture recognition. In: 2018 24th

international conference on pattern recognition (ICPR). IEEE, 2018. p. 3451-3456.
32. NUNEZ, Juan C., et al. Convolutional neural networks and long short-term memory for skeleton-based human activity and hand gesture recognition. Pattern Recognition, 2018, 76: 80-94.
33. Chung H., Chung Y., Tsai W. An efficient hand gesture recognition system based on deep CNN; Proceedings of the 2019 IEEE International Conference on Industrial Technology (ICIT); Melbourne, Australia. 13–15 February 2019; pp. 853–858
34. Li G., Tang H., Sun Y., Kong J., Jiang G., Jiang D., Tao B., Xu S., Liu H. Hand gesture recognition based on convolution neural network. Cluster Comput. 2019;22:2719–2729. doi: 10.1007/s10586-017-1435
35. HTAY, Than Than; MAUNG, Su Su. Early stage breast cancer detection system using glcm feature extraction and k-nearest neighbor (k-NN) on mammography image. In: *2018 18th International Symposium on Communications and Information Technologies (ISCIT)*. IEEE, 2018. p. 171-175.
36. QURRATU'AINI, Dayang, et al. Visual-based fingertip detection for hand rehabilitation. *Indonesian Journal of Electrical Engineering and Computer Science*, 2018, 9.2: 474-480.
37. Hernández, Irene. Automatic Irish sign language recognition. University of Dublin, Trinity College, 2018.
38. RATHI, Dhruv. Optimization of transfer learning for sign language recognition targeting mobile platform. *arXiv preprint arXiv:1805.06618*, 2018.
39. Daroya, Rangel & Peralta, Daryl & Naval, Prospero. (2018). Alphabet Sign Language Image Classification Using Deep Learning. 0646-0650. 10.1109/TENCON.2018.8650241.
40. JIANG, Xianwei; ZHANG, Yu-Dong. Chinese sign language fingerspelling via six-layer convolutional neural network with leaky rectified linear units for therapy and rehabilitation. Journal of Medical Imaging and Health Informatics, 2019, 9.9: 2031-2090.
41. BUSTAMAM, A., et al. Application of K-means clustering algorithm in grouping the DNA sequences of hepatitis B virus (HBV). In: *AIP Conference Proceedings*. AIP Publishing LLC, 2018. p. 030134.

42. MOHAMMED, Bakhtyar Ahmed; AL-ANI, Muzhir Shaban. Digital medical image segmentation using fuzzy C-means clustering. UHD Journal of Science and Technology, 2020, 4.1: 51-58.
43. LIU, Yanqiu; WANG, Xiuhui; YAN, Ke. Hand gesture recognition based on concentric circular scan lines and weighted K-nearest neighbor algorithm. *Multimedia Tools and Applications*, 2018, 77.1: 209-223.
44. YANG, Pei, et al. An improved Otsu threshold segmentation algorithm. *International Journal of Computational Science and Engineering*, 2020, 22.1: 146-153.
45. PRABHU CHAKKARAVARTHY, A.; CHANDRASEKAR, A. An Automatic Threshold Segmentation and Mining Optimum Credential Features by Using HSV Model. *3D Research*, 2019, 10.2: 1-17.
46. SEKEHRAVANI, Ehsan Akbari; BABULAK, Eduard; MASOODI, Mehdi. Implementing canny edge detection algorithm for noisy image. Bulletin of Electrical Engineering and Informatics, 2020, 9.4: 1404-1410.
47. Aithal, C. N., Ishwarya, P., Sneha, S., Yashvardhan, C. N., Kumar, D., & Suresh, K. V. (2021, November). Dynamic Hand Segmentation. In 2021 IEEE 8th Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON) (pp. 1-6). IEEE. https://doi.org/10.1109/UPCON52273.2021.9667634.
48. Son Nguyen, X., Brun, L., Lézoray, O., & Bougleux, S. (2019). A neural network based on SPD manifold learning for skeleton-based hand gesture recognition. arXiv e-prints, arXiv-1904. https://doi.org/10.48550/arXiv.1904.12970.
49. HALDER, Arpita; TAYADE, Akshit. Real-time vernacular sign language recognition using mediapipe and machine learning. Journal homepage: www. ijrpr. com ISSN, 2021, 2582: 7421
50. Chua, S. N., Chin, K. Y., Lim, S. F., & Jain, P. (2022). Hand Gesture Control for Human–Computer Interaction with Deep Learning. Journal of Electrical Engineering & Technology, 17(3), 1961-1970.EDDY. https://doi.org/10.1007/s42835-021-00972-6 .
51. Zhang, Fan et al. “MediaPipe Hands: On-device Real-time Hand Tracking.” *ArXiv* abs/2006.10214 (2020).

52. KROTOV, Dmitry; HOPFIELD, John J. Dense associative memory for pattern recognition. *Advances in neural information processing systems*, 2016, 29.
53. KAREEM, Emad I. Abdul; ALSALIHY, Wafaa AH Ali; JANTAN, Aman. Multi-Connect Architecture (MCA) associative memory: A modified hopfield neural network. *Intelligent Automation & Soft Computing*, 2012, 18.3: 279-296.
54. https://www.tutorialspoint.com/artificial_neural_network/artificial_neural_network_associate_memory.htm.15/8/2023
55. He, H., Shang, Y., Yang, X., Di, Y., Lin, J., Zhu, Y. et al Constructing an associative memory system using spiking neural network. Frontiers in neuroscience, 2019, 13: 650. https://doi.org/10.3389/fnins.2019.00650.
56. KAREEM, Emad I. Abdul; ABBAS, Safana H.; SALMAN, Salman Mahmood. Intelligent traffic light controller based on MCA associative memory. *Science Journal of Circuits, Systems and Signal Processing*, 2014, 6.6-1: 6-16.
57. HAGHIGHI, Sepand, et al. PyCM: Multiclass confusion matrix library in Python. *Journal of Open Source Software*, 2018, 3.25: 729.
58. Elatawy, S. M., Hawa, D. M., Ewees, A. A., & Saad, A. M. Recognition system for alphabet Arabic sign language using neutrosophic and fuzzy c-means. Education and Information Technologies, 2020, 25.6: 5601-5616.
59. DAS, Prangon; AHMED, Tanvir; ALI, Md Firoj. Static hand gesture recognition for American Sign Language using deep convolutional neural network. In: 2020 IEEE Region 10 Symposium (TENSYMP). IEEE, 2020. p. 1762-1765.https://doi.org/10.1109/TENSYMP50017.2020.9230772.
60. Abdul, W., Alsulaiman, M., Amin, S. U., Faisal, M., Muhammad, G., Albogamy, F. R., & Ghaleb, H. (2021). Intelligent real-time Arabic sign language classification using attention-based inception and BiLSTM. Computers and Electrical Engineering, 95, 107395. https://doi.org/10.1016/j.compeleceng.2021.107395
61. DE CASTRO, Giulia Zanon; GUERRA, Rúbia Reis; GUIMARÃES, Frederico Gadelha. Automatic translation of sign language with multi-stream 3D CNN and generation of artificial depth maps. Expert Systems with Applications, 2023, 215: 119394.

المستخلص

التعرف على إيماءات اليد تحت رؤية الكمبيوتر أمرًا بالغ الأهمية لتسهيل التواصل بين مجتمع الصم والبكم والأشخاص العاديين، وكذلك بين كبار السن الذين لا يستطيعون ارتداء قفازات اليد أو أجهزة الاستشعار. يعد تفسير إيماءات اليد جزءًا أساسيًا من التفاعل بين الإنسان والحاسوب والذي يعتمد بشكل كبير على القدرة على فهم الإشارات البشرية والاستجابة لها.

يعد التعرف على إيماءات اليد المعقدة للغاية والمتشابهة أمرًا بالغ الأهمية. ونتيجة لذلك، هناك حاجة ملحة لنظام يتمتع بالدقة والقدرة على التعرف على هذه الإيماءات، حيث أن مشكلة التعرف بشكل غير دقيق على إيماءات اليد قد تضر بالمجتمع البشري الذي يعتمد على الإيماءات لنقل رغباته. نظرًا لأهمية تطبيقات التعرف على إيماءات اليد، فقد تم تطوير طريقة جديدة تمامًا للتعرف على إيماءات اليد وتطبيقها على مجموعة بيانات إيماءات اليد.

هدف النظام المقترح هو إلقاء الضوء على الخطوات الأكثر أهمية في عملية الكشف عن إيماءات اليد. وهي عملية الكشف والتعرف على إيماءات اليد باستخدام نقاط معالم اليد للكشف عن الأيدي واستخدام بنية Multi-Connect Architecture MCA للذاكرة الترابطية كاتجاه جديد في مرحلة التعرف. وحل مشكلة التشابه بين العلامات بسبب الارتباط القوي بين حركات الأصابع. بالإضافة إلى عدم الدقة العالية بين العلامات المعقدة والمتشابهة جداً مثل (A ، S، E) ومشكلة زمن الاستجابة في التعرف على إيماءات اليد في الوقت الحقيقي، وبالتالي فإن استخدام الشبكة العصبية للذاكرة الترابطية MCA أدى إلى تحسين الكفاءة في التعامل مع الارتباط بين الأنماط المتشابهة من خلال أخذ متجهات مماثلة لكل نمط من إيماءات اليد مرة واحدة فقط.

أعطى النظام المقترح دقة عالية في الوقت الحقيقي من خلال استخدام مصفوفة الارتباك وأظهر نتائج واعدة بدقة لغة الإشارة الأمريكية والصينية والعربية والأرقام بنسبة 95.42%، 92.13%، 93.55%، و94% على التوالي. وكذلك عمل النظام في بيئة غير خاضعة للرقابة، بالإضافة إلى كونه نظام قابل للتطبيق من خلال تحويل الإشارة إلى معناها وجمعها ككلمات وجمل وليس مجرد حروف.

وزارة التعليم العالي والبحث العلمي

جامعة بابل

كلية تكنولوجيا المعلومات

قسم البرمجيات

التعرف على إيماءات اليد استنادًا إلى الذاكرة الترابطية للهندسة المعمارية متعددة التوصيلات

أطروحة مقدمة الى مجلس كلية تكنولوجيا المعلومات في جامعة بابل

وهي جزء من متطلبات الحصول على درجة الدكتوراه فلسفة في

تكنولوجيا المعلومات/ البرمجيات

من قبل

نور فاضل حسين سليمان

بإشراف

الاستاذ الدكتور

عماد عيسى عبد الكريم

2023م

1445 هـ