

Original Research Paper

Arabic Static and Dynamic Gestures Recognition Using Leap Motion

Basma Hisham and Alaa Hamouda

Department of Computer Engineering, Al-Azhar University, Cairo, Egypt

Article history

Received: 15-05-2017

Revised: 07-07-2017

Accepted: 11-08-2017

Corresponding Author:
Basma Hisham
Department of Computer
Engineering, Al-Azhar
University, Cairo, Egypt
Email: basmahisham.2015@gmail.com

Abstract: Across the world, several millions of people use sign language as their main way of communication with their society, daily they face a lot of obstacles with their families, teachers, neighbours, employers. According to the most recent statistics of World Health Organization, there are 360 million persons in the world with disabling hearing loss i.e. (5.3% of the world's population), around 13 million in the Middle East. Hence, the development of automated systems capable of translating sign languages into words and sentences becomes a necessity. We propose a model to recognize both of static gestures like numbers, letters, ...etc and dynamic gestures which includes movement and motion in performing the signs. Additionally, we propose a segmentation method in order to segment a sequence of continuous signs in real time based on tracking the palm velocity and this is useful in translating not only pre-segmented signs but also continuous sentences. We use an affordable and compact device called Leap Motion controller, which detects and tracks the hands' and fingers' motion and position in an accurate manner. The proposed model applies several machine learning algorithms as Support Vector Machine (SVM), K- Nearest Neighbour (KNN), Artificial Neural Network (ANN) and Dynamic Time Wrapping (DTW) depending on two different features sets. This research will increase the chance for the Arabic hearing-impaired and deaf persons to communicate easily using Arabic Sign language(ArSLR). The proposed model works as an interface between hearing-impaired and normal persons who are not familiar with Arabic sign language, overcomes the gap between them and it is also valuable for social respect. The proposed model is applied on Arabic signs with 38 static gestures (28 letters, numbers (1:10) and 16 static words) and 20 dynamic gestures. Features selection process is maintained and we get two different features sets. For static gestures, KNN model dominates other models for both of palm features set and bone features set with accuracy 99 and 98% respectively. For dynamic gestures, DTW model dominates other models for both palm features set and bone features set with accuracy 97.4% and 96.4% respectively.

Keywords: Sign Language, Leap Motion Controller, Static Gestures, Dynamic Gestures

Introduction

Sign language is the most common and important way for deaf and hearing impaired in order to communicate and integrate with their society. It is a kind of visual language that consists of a sequence of grammatically structured human gestures (Quesada *et al.*, 2015). There is a large sector of Arabian community suffering from deafness and hearing impaired. In Egypt, the number of

deaf people according to "Central Agency for Public Mobilization and Statistics" last study is around 2 million and increased in 2012 to be close to 4 million (<http://www.who.int/mediacentre/factsheets/fs300/en/>). Unfortunately, most of these people cannot read or write Arabic language and 80% of them are literal, they are isolated from their society. They are a large part of society and cannot be neglected however still they cannot communicate normally with their community

because of the constraints of language known to them, most people are not familiar with sign language and cannot understand it. Thus, they are far from their own society, depressed, living lonely life. So, these restrictions must be broken because they prevent the deaf persons from enjoying his full rights and get their opportunities for full citizenship. For example, the deaf person must be able to express himself and what he wants easily. Hence, it is important to develop automated system capable of translating sign languages into words and sentences. This will help normal people to communicate effectively with the deaf and the hearing-impaired and will act as an interface between normal person who does not know the sign language and the deaf person.

The sign recognition approached can be categorized as sensor-based approach and image-based approach. In sensor-based approach the deaf person needs to wear external instruments such as electronic gloves which contain number of sensors during performing the signs to detect the different hands and fingers motions. In the Image-based systems the camera(s) are used to acquire the images of the hand during the motion (Samir Elons *et al.*, 2013). However, the two approaches have advantages and disadvantages such as, in the sensor-based the data acquisition and data processing are simple and its reading and results are accurate, but in this approach the user is enforced to wear external gloves and this leads to difficulty in interactions and inflexibility in movements (Mohandes *et al.*, 2014). The image processing approach allow the natural interaction for the user but it needs specific background and environmental conditions also it needs specific light intensity to achieve high accuracy also it needs complex computational in order to process and analysis the captured images (Vijay *et al.*, 2012).

In this study, a new model for Arabic Sign Language Recognition (ArSLR) was developed for both static and dynamic gestures. It was built using the recently introduced sensor called Leap Motion Controller (LMC). Leap Motion Controller (LMC) was developed by an American company called leap motion company. It detects and tracks position and motion for hands, fingers and their joints with a rate of 200 frames per second approximately. The captured frames contain information about how many hands are being detected and also vectors that contain information about the position and rotation of hands and fingers based a skeletal model of the hand. The proposed models depend on two different sets of features called Palm-Features set and Bone-Features set which have common features between them. Several experiments are performed to test the proposed models, the experiments include both static gestures and dynamic gestures i.e., Arabic alphabets, Arabic numerical and the common signs used with Dentist for static gestures and common verbs and nouns using one hand and two hands for dynamic gestures. The static gestures include fixed gestures which haven't any type of motions or movement

but the dynamic gestures contain movement and motion in performing it. Machine learning algorithms like Support Vector Machine (SVM), K-Nearest Neighbor (KNN), Artificial Neural Network (ANN) and Dynamic Time Wrapping (DTW) are used to classify the datasets. Also, the system develops a method to segment a sequence of signs in order to recognize the continuous sentences which makes the system more reliable and more practical.

The main aim of this paper is to propose a recognition model for Arabic sign language using Leap Motion Controller (LMC) and apply it on static and dynamic gestures by choosing the optimal features with acceptable accuracy. The rest of this paper is organized as follows: In section 2, related work of Arabic sign language recognition is presented. Introduction of Leap Motion Controller is presented in section 3. The methodology and experimental results are presented in section 4 and 5 respectively. The segmentation part is presented in section 5. Finally, conclusion and future work is presented in section 6.

Related Work

The research on Arabic Sign Language Recognition (ArSLR) is recent compared with the work carried out in the field of other sign languages recognition. There are several researches and techniques have been developed but there is a big challenge in this field and it is still an open area for further research especially ArSL because most of sign languages have data bases and dictionaries which help the researchers in their works but ArSLR researches are still recent and no such data sets. Mohandes (2001) developed a model to recognize Arabic sign language alphabets, the support vector machine is used for classification, it was fed by the moment invariants and the features extracted using Hus moments, the recognition accuracy was 87%. Al-Jarrah and Halawani (2001), proposed a translator for 30 Arabic manual alphabets based on a neurofuzzy inference system, the system depends on image based approach and the captured images are segmented, processed, analyzed and converted to a set of features. The system achieved accuracy of 93.55%. Assaleh and Al-Rousan (2005), the system depends on sensor -based approach, the signer should use a glove with 6 different colors, 5 of them for fingertips and one for the wrist region, the model uses a polynomial classifier to recognize alphabet signs. A recognition rate was around 93.4%, the evaluation was done on a database of more than 200 samples for 42 gestures. In Mohandes and Buraiky (2007) used cost effective instrumented gloves to implement a robust and accurate ArSLR system. The Statistical features are extracted from the acquired signals and the gestured were classified using SVM classifier. The model was evaluated and tested on a database of 120 signs, the recognition accuracy was over 90%. In Maraqa and Abu-Zaiter (2008) built a

model for alphabets recognition using recurrent neural networks. This model covered 30 gestures and the training dataset was 900 samples collected from two users. Colored gloves were used in their experiment. The recognition rate that has reached up to 95.11% In El-Bendary *et al.* (2010) presents an automatic translation system of gestures of the manual alphabets in the Arabic sign language. The system depends on Image-based approach, the main steps in this system are; pre-processing phase, frame detection phase, category detection phase, features extraction phase and classification phase. The single nearest neighbor technique is used for classification. Experiments results of this system were able to recognize the 30 Arabic alphabets with an accuracy of 91.3%. Hemayed and Hassanien (2010) introduced a new model for hand gesture recognition to recognize Arabic sign language alphabet and finally converts it into voice correspondences, it was based on Image-based approach. The input was the color image which was captured for the hand motion, then it was converted to YCbCr color space and then extract the skin region from colored images. Prewitt edge detector is used to extract the edges of the segmented hand gesture. In classification phase Principal Component Analysis (PCA) is used with a K-Nearest Neighbor Algorithm (KNN). They applied the technique on more than 150 signs and their accuracy was near to 97% at real time test for three different signers. In Shanableh and Assaleh (2011) presented a system for independent Arabic sign recognition using a video-based approach. The test experiment includes 3450 video segments covering 23 isolated gestures from three signers. The signers used colored gloves and the color information was used in the preprocessing phase. The input was image sequence through successive image differencing and the output was extracted features that were used to detect the motion information. The used classifier was KNN and the achieved recognition rate was 87%. In Mohandes *et al.* (2012) developed a system for Arabic signs Recognition, it was a vision-based model which used a Hidden Markov Model (HMM) to identify the pre-segmented Arabic signs from images. The experiments were done on a dataset consisting of 500 samples of 300 signs and achieved a recognition accuracy of 95%. In Mohandes (2013), two-handed Arabic sign recognition was introduced. The database which was used for evaluation consisted of 20 samples from each of 100 two-handed signs performed by two signers. For classification, the SVM was used, achieving an accuracy of 99.6% with 100 signs. In Guesmi *et al.* (2016) presented an automatic system for Arabic sign language recognition, they recognized in real time the static hand gestures of the Arabic sign language and then convert them into Arabic text. There were two main phases in the system: Hand detection and hand gesture recognition. The

system using two classifiers Fast Wavelet transform (FWNC) and the Separator Wavelet Network Classifier (SWNC). The hand detection experiment contained 100 images, the results of (FWNC) was 99.20% and (SWNC) was 92.36%. The hand recognition experiment contained 28 signs for Arabic alphabet, the results of (FWNC) was 93.21% and (SWNC) was 71.07%.

Aly *et al.* (2016) developed a new system for Arabic finger spelling recognition. The system is a type of image-based approach. The input to the system is the two images which are captured from Softkinect camera. The information that were extracted from the images such as color and depth of the images are used to classify each hand spell. The accuracy achieved was 99.5%.

In addition to the different image-based and glove-based systems that are currently in use, some new systems for facilitating human-machine interaction have been introduced lately. Microsoft Kinect and the Leap Motion Controller (LMC) have attracted special attention. The Microsoft Kinect system uses an infrared emitter and depth sensors, in addition to a high-resolution video camera.

In 2011, depth information from Kinect sensor is used by Biswas and Basu (2011) recognize signs in Japanese Sign Language (JSL). They use low level features and achieve more than 90% accuracy on 8 signs, Aliyu *et al.* (2016) proposed an Arabic sign language recognition system based on Microsoft Kinect system. The developed system was tested with 20 signs from the Arabic sign language dictionary. As stated before, the sensor based approach enforced the user to wear external hardware equipment (i.e., Gloves) and this is difficult in interactions and inflexible in movements, also there is another issue related to calibration because different people have different hand sizes and finger length/thickness. As well, image-based approach has challenge of environmental conditions such as lighting conditions, image background and different types of noise. Microsoft Kinect gave good results and does not enforce the user to wear any external hardware equipment. As well, the collected data is independent of the environmental conditions. But it can't detect the details of hand and fingers. So it fails in recognizing the gestures that are performed by hands but leap motion controller is very accurate in hand tracking and its details. So, we use it in this research for hand gesture recognition.

The Leap Motion Controller (LMC)

Recently, new systems and devices appeared for enabling human machine interaction to be easier. Such as, Leap Motion Controller (LMC) which was developed by an American company called leap motion company (<http://dartmouthbusinessjournal.com/2013/08/the-leap->

motion-controller-and-touchless-technology). It detects and tracks position and motion for hands, fingers and their joints. It operates at a rate of 200 frames per second approximately. These frames contain information about how many hands/tools are being detected and also vectors that contain information about the position and rotation of hands and fingers based a skeletal model of the hand that is represented in Fig. 1. The effective range of the LMC extends from 1 inch to 2 feet above the device (<http://phys.org/news/2013-08-motion-ready-prime.html>). The LMC uses two high precision infrared cameras and three Light-Emitting Diodes (LEDs) to capture hand information within its active range (<https://www.leapmotion.com>) it is precise and handy in every day situation and also low cost. Figure 2 represents leap motion controller. There are several researches in ARSL which uses leap motion controller and the results were promising (Mohandes *et al.*, 2014; Elons *et al.*, 2014).

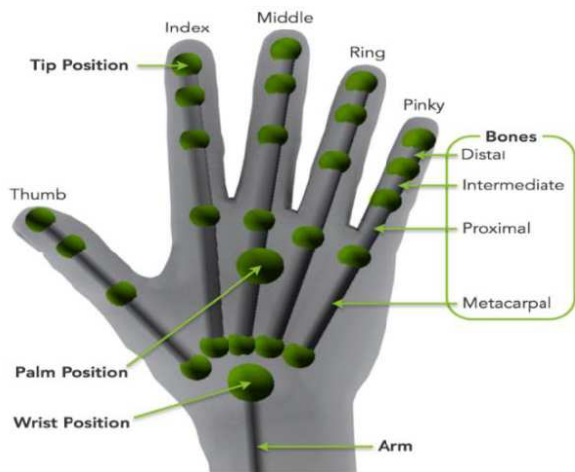


Fig. 1. Leap motion Hand Tracking (<http://blog.leapmotion.com/getting-started-leap-motion-sdk/>)

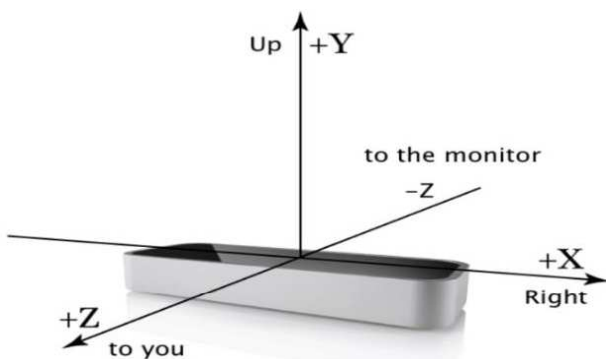


Fig. 2. Coordinate system of Leap Motion Controller (<http://developer-archive.leapmotion.com/getting-started/javascript/developer-guide>)

Methodology

The overall workflow of the proposed model is composed of five steps as shown in Fig. 3. The first step is pre-processing phase which starts up the leap motion service in order to receive the hand gestures as input and perform some advanced algorithms on the received raw sensory data. The second step is tracking phase, in this phase the tracking layer matches the data to extract tracking information such as, fingers and tools. The third step concerns with the features extraction, the data obtained from LMC are analyzed to extract robust features that can be used to identify the signs these features are introduced as a vector to a classifier and used to identify the signs. Finally, the last step is to use the trained classifier to recognize users' gesture.

Preprocessing

The Leap Motion Service is the software on the computer that makes set of processes and mathematical operations on the received images. The preprocessing phase consists of several steps such as background compensation for objects (for example heads) and another environmental factor as lighting, then perform infrared scanning for the received images to construct a 3D representation for what the device sees as hands and tools, finally create a digital version of the hand in areal time.

Tracking

In this phase, some algorithms and mathematical operations are used to interpret the 3D data and infer the positions of detected objects, finally return the results which are expressed as a series of frames which contain information about the hand motion, the position of hands, fingers and their joints in 3D coordinates.

Features Extraction

The used features in the model were Palm-Features set and Bone-Features set, there are some common features between the two sets such as Palm position in the three directions (x, y, z), Palm Direction in (x, y, z) also Fingertips direction in (x, y, z) for every finger (Thumb, Index, Middle, Rin, Pinky) (Fig. 4) Pitch, Yaw and Roll angles (Fig. 5), The vector that resulted from the angles between the fingers (Fig. 6).

The palm vector which contains six scalar values and it will be as the following:

$$P_set = \left\{ \begin{array}{l} P_Pos_x, P_Pos_y, P_Pos_z, \\ P_Dir_x, P_Dir_y, P_Dir_z \end{array} \right\}$$

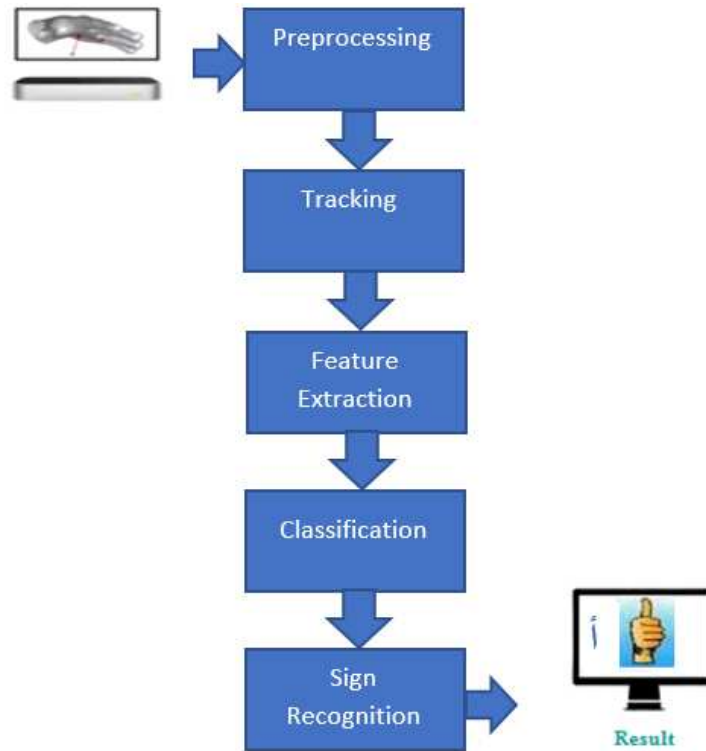


Fig. 3. Proposed model workflow



Fig. 4. Finger tip direction

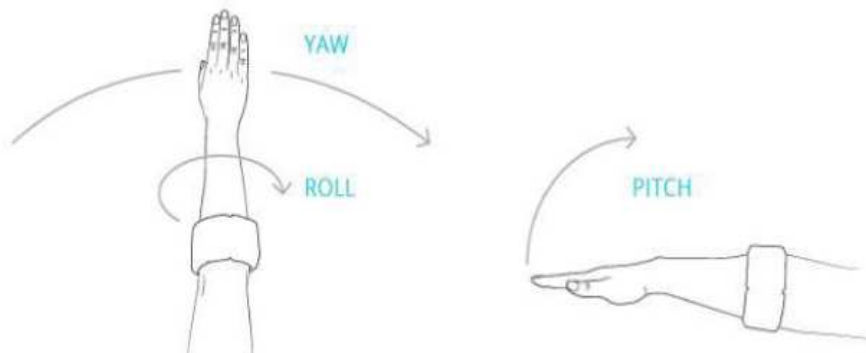


Fig. 5. Pitch, yaw and roll

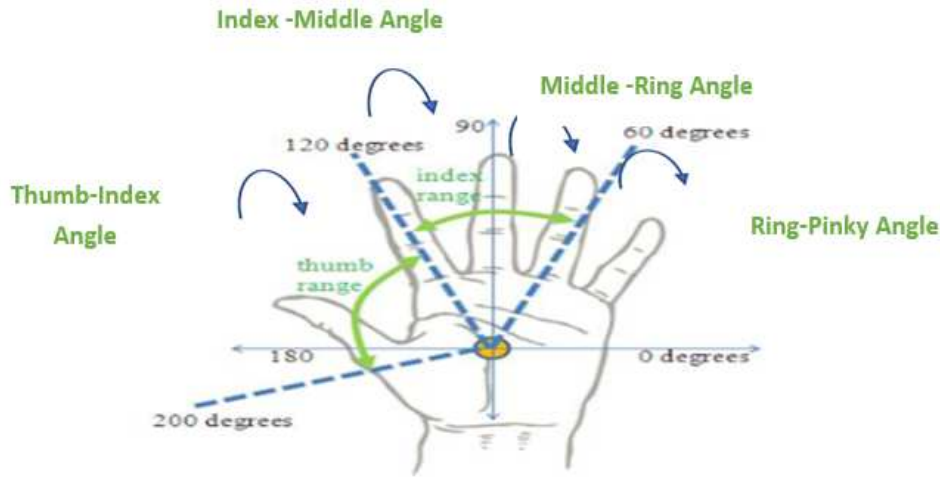


Fig. 6. Angles between fingers

The fingertips direction vector which contains fifteen scalar values and it will be as the following:

$$F_Tip_set = \{d_1, d_2, d_3, d_4, d_5\}$$

where, d_i is the direction of finger that is indexed with i and $i \in \{\text{Thumb, Index, Middle, Ring, Pinky}\}$.

The Arm_Dir_set direction vector which contains three values and it will be as the following:

$$Arm_Dir_set = \{Pitch_{Angle}, Yaw_{Angle}, Roll_{Angle}\}$$

The $Finger_Angles_set$ direction vector which contains three scalar values and will be as the following:

$$Finger_Angles_set = \{Thumb_Index_{Angle}, Index_Middle_Ring_{Angle}, Ring_Pinky_{Angle}\}$$

For the Palm features set, in addition to the previous features there are another vector resulted from the subtraction of each phalanx's end position and the Palm's position P as a reference point, each value was represented in three coordinated (x, y, z) and this vector contains 57 scalar values (Fig. 7):

$$H_set = \left\{ \begin{array}{l} F_{11} - P, F_{12} - P, F_{13} - P, \\ F_{14} - P, \dots, F_{52} - P, F_{53} - P \end{array} \right\}$$

This vector along with previous vectors P_set vector for palm position and palm direction, F_Tip_set for fingertips direction, Arm_Dir_set for arm direction angels (Pitch, Yaw, Roll) and $Finger_Angles_set$ for angles between fingers the combination of all of these vectors contains 85 scalar values and represented the features set S1:

$$S1 = \left\{ \begin{array}{l} H_set, P_set, F_Tip_set, \\ Arm_Dir_set, Finger_Angles_set \end{array} \right\}$$

For the Bone features set, also in addition to the previous features P_set vector for palm position and palm direction, F_Tip_set for fingertips direction, Arm_Dir_set for arm direction angels (Pitch, Yaw, Roll) and $Finger_Angles_set$ for angles between fingers, there is a vector resulted from subtracting the phalanx's start position from phalanx's end position for every finger the resulted value in (x, y, z) this vector contains 42 scalar values (Fig. 8):

$$H_set = \{F_{11} - F_{12}, F_{13} - F_{12}, \dots, F_{24} - F_{23}, F_{53} - F_{52}\}$$

This vector along with previous vectors P_set vector for palm position and palm direction, F_Tip_set for fingertips direction, Arm_Dir_set for arm direction angels (Pitch, Yaw, Roll) and $Finger_Angles_set$ for angles between fingers the combination of all of these vectors contains 70 scalar values and represented the feature set S2:

$$S2 = \left\{ \begin{array}{l} H_set, P_set, F_Tip_set, \\ Arm_Dir_set, Finger_Angles_set \end{array} \right\}$$

Single-Sign Classification

There are two types of gesture: Static and dynamic gestures. The group of static gestures includes fixed gestures, which are not taking into account the changes in time such as alphabets, numbers also pointing gestures (Fig. 9) that are including pointing to spatial location or specific object, static semaphoric gestures (Fig. 10) which are type of static gestures such as thumbs-up meaning approval symbol and iconic gestures (Fig. 11) which used to represent shape, size, curvature of object or entities (Benko *et al.*, 2012).

The dynamic gestures contain movements and motion during performing it, such as pointing and Manipulation gestures which are used to guide movement in a short feedback loop (Fig. 12) and dynamic iconic gestures which are often used to describe paths or shapes, such as moving the hand in circles, meaning "the circle" (Fig. 13).

Static single sign Classification

So, for static single sign, the classifier takes its input as a single frame which represents the sign in, we tested several classifiers and compare their results such as Support Vector Machine (SVM) classifier which is one of the most common machine learning classifiers and was widely used in object detection and recognition also (KNN) and (ANN), these methods are chosen because it took into consideration the state-of-the-art methods for many different applications and they gave the best results.

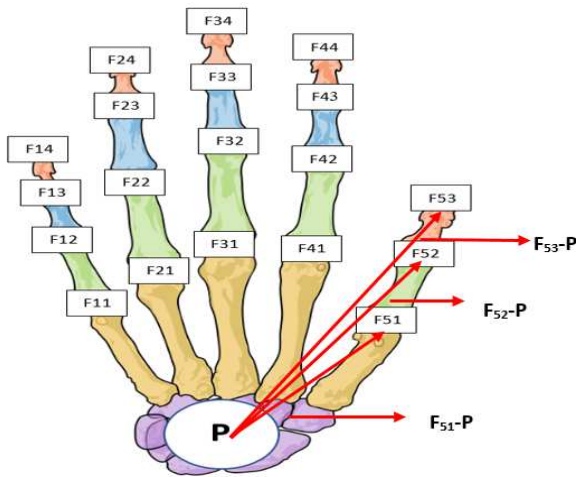


Fig. 7. Palm-features set

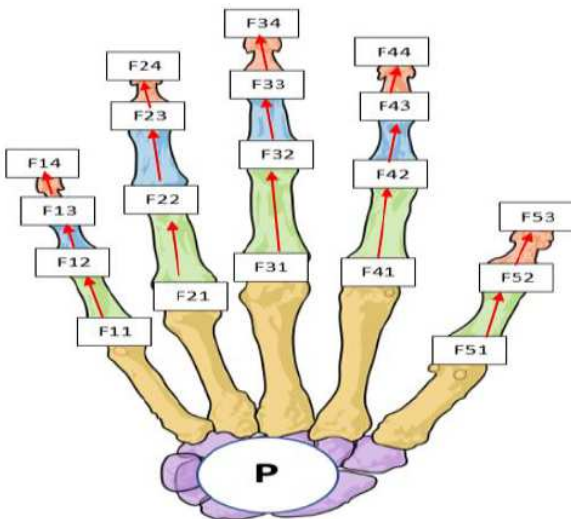


Fig. 8. Bone-feature set



Fig. 9. Pointing gesture

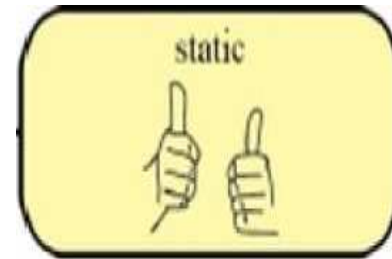


Fig. 10. Static semaphoric gesture

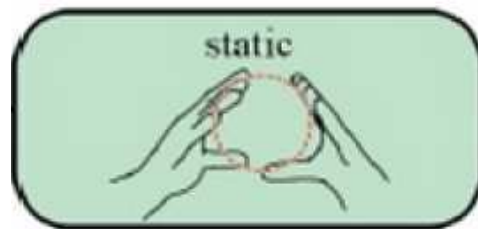


Fig. 11. Static iconic gestures

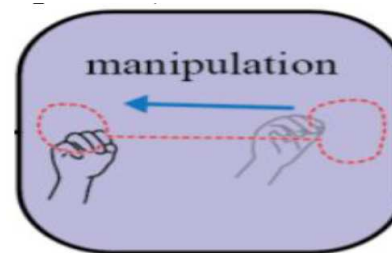


Fig. 12. Dynamic manipulation gestures

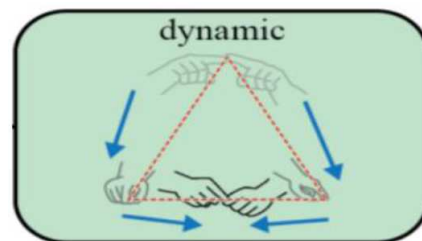


Fig. 13. Dynamic iconic gestures

Dynamic Single Sign Classification

The second type of signs is the dynamic single sign which contains movement and motion in performing it. For this type, there are two suggested methods.

In the first method, the classifier took a sequence of frames that compose a single sign and we used the classifiers as SVM or KNN or ANN and then classify each frame individually and the result is selected by simple majority (i.e.,) the classification with the most frames assigned to it (Fig. 14).

In the second method, Dynamic Time Warping (DTW) was used, it is known as an optimal alignment algorithm between two given sequences it measures the similarity between two sequences which are varying in time or speed, DTW has been used in various fields, such as speech recognition data mining and movement recognition. The

DTW is very suitable for sign recognition applications, because of coping in that way with sign executions speeds (https://en.wikipedia.org/wiki/Dynamic_time_warping). In this case the set of frames in Test (The Test Sign) set will be compared with a set of frames in training set using DTW and each set will be treated as a signal or pattern. The sequences that will be compared must be "warped" non-linearly in the time dimension to determine their similarity independent of certain non-linear variations in the time dimension. The DTW will determine the most similar group in the training set to the test sign according to the calculated distance, so the most similar sequence to the test sequence is the sequence in the training set with the smallest DTW distances. Finally, it gets the most similar group and assign it to the test sign. The model of Dynamic sign recognition is represented in the (Fig. 15).

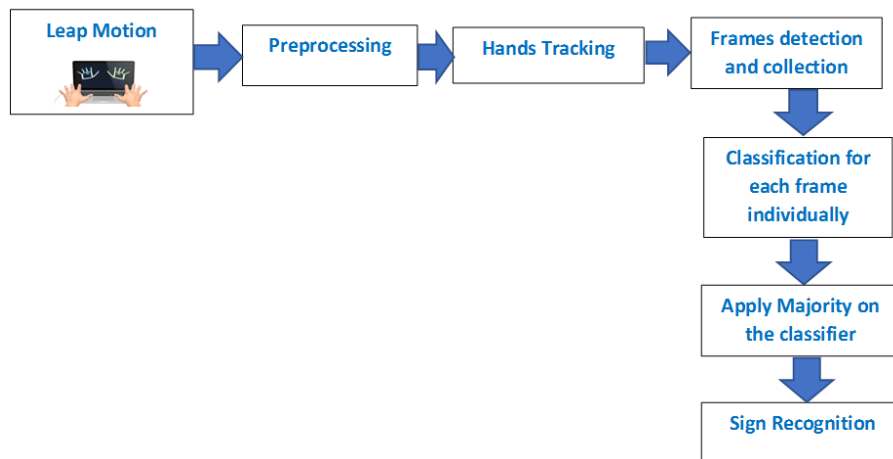


Fig. 14. Static single sign classification

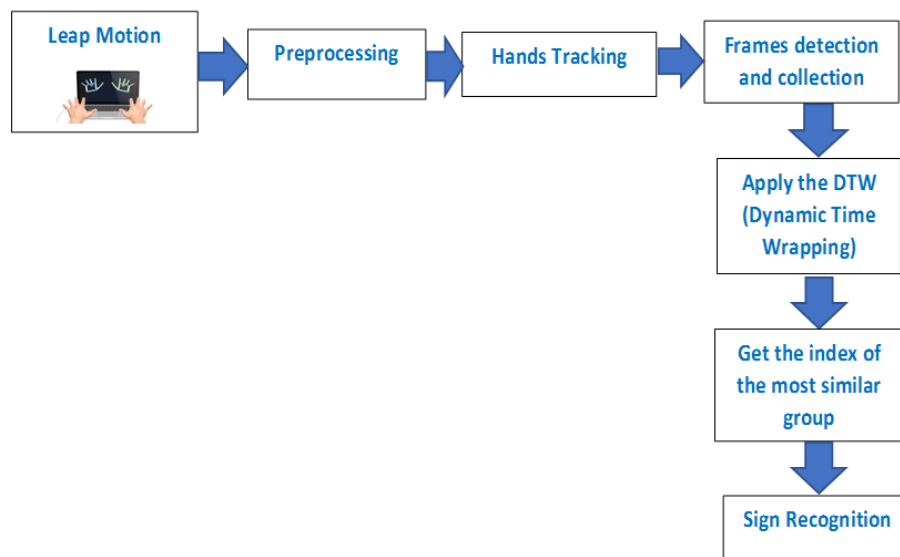


Fig. 15. Dynamic single sign classification

Experiments Results

We performed several experiments to test the developed models, the experiments include Arabic alphabets, Arabic numerical, the common signs used with Dentist, the common verbs and nouns using one hand and, common verbs and Nouns using two hands.

Arabic Alphabets

The developed system is used to recognize the twenty-eight Arabic alphabet signs from أ-ى as shown in Fig. 16. It should be noted that all Arabic alphabet signs are type of static gestures and are performed using a single hand. The training set was about 400 frames for each sign (Alphabet) and collected from two different users and tested by 200 frames from another third user, the number of features are 85 for palm features set and 70 features for bone features set.

The results of classifications using KNN, SVM and ANN are shown below in Fig. 17.

Arabic Numbers

Our developed system also is used to recognize the Arabic signs that represent the numbers from 0 to 10

which are shown in Fig. 18. It is also a type of static one hand gestures. The training set was about 400 frames for each sign (Number) collected from two different users and tested by 200 frames from another third user.

The results of classifications using KNN, SVM and ANN are shown below in Fig. 19.

Common Dentist Signs

The group of common dentist signs are consisting of eight words some of them are static gestures such as: انا (It is Ok), يؤلم قليلا (Little painful), يؤلم كثيرا (Very painful), توقف (Stop), ألم (Pain). Other Signs are dynamic such as اريد شطف السوائل (Want to Rinse the fluids), اريد التحدث (Want talking), اريد حك وجهي (Want to knit the face) but both are performed by a single hand, they are shown in Fig. 20.

The equivalent English meaning are represented in Table 1.

For Static Gestures, the training data set in classification method contains 50 samples for each sign from two different users and tested by 20 samples from another user, the comparisons between the used classifiers (KNN, SVM and ANN) in static dentist signs are shown in Fig. 21.



Fig. 16. Arabic alphabet signs

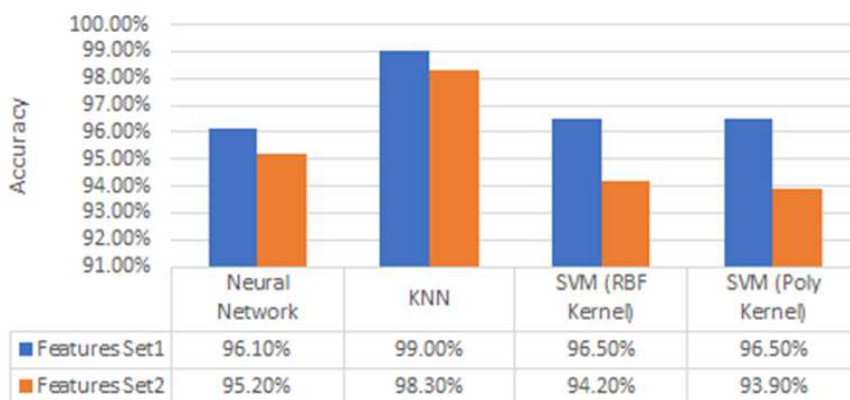


Fig. 17. Arabic alphabet signs features comparison

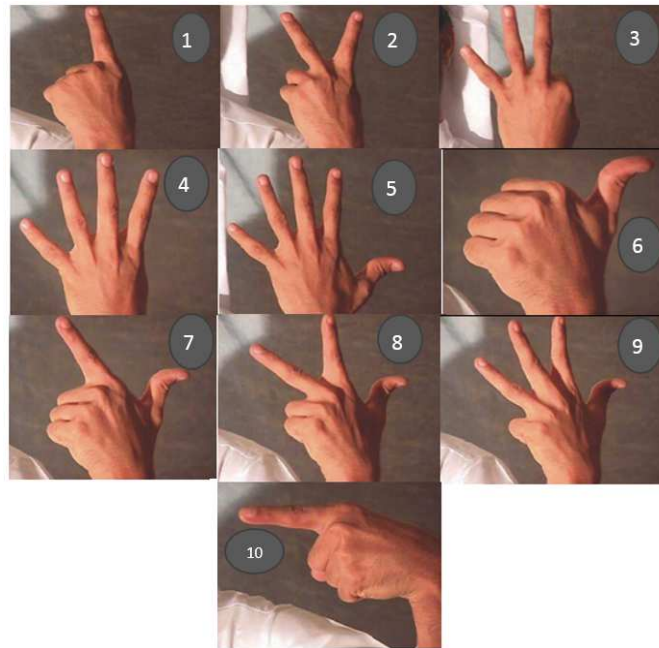


Fig. 18. Arabic numbers signs

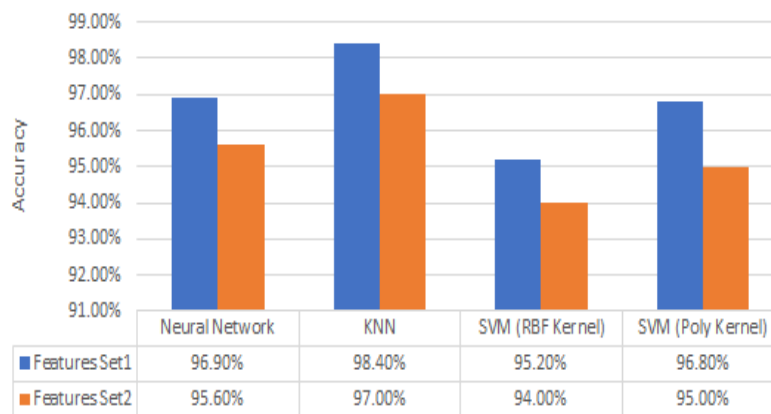


Fig. 19. Arabic numbers signs features comparison

Table 1. English meaning for common dentist signs

Arabic sign	Meaning in English
الم	Pain
انا بخير	It is Ok
توقف	Stop
اريد التحدث	Want Talking
يؤلم قليلا	Little painful
يؤلم كثيرا	Very Painful
اريد حك وجهي	Want to knit the face
اريد شطف السوائل	Want to Rinse the fluids

As I explained before that the dynamic gestures are represented by a set of frames, so for recognition of this type we suggested "Classification + Majority" and DTW (Dynamic Time Wrapping). The comparisons between the used classifiers (KNN, SVM and ANN) + Majority in dynamic dentist signs are shown in Fig. 22.

The result of the second method "DTW" is listed in Table 2.

Common Verbs and Nouns with One Hand

The group of this experiment contains common words and verbs. The number of these signs are twenty words, some of them are static gestures such as the following (هاتف (Phone), زواج (Marriage), نجاح (Success), حب (Love), اب (Father), ام (Mother), جد (Grand Father), طفل (Children)) and some are dynamic such as the following:

(خال ((Maternal)uncle), اخت (Sister), اخ (Brother), تعال (Come), عم (Uncle), عم (Aunt), عم (Grand Mother), عم (Uncle), عم (Aunt), عم (Grand Mother)) but both are performed by a single hand, they are shown in Fig. 23., also the equivalent English meaning are shown in Table 3.



Fig. 20. Common dentist signs

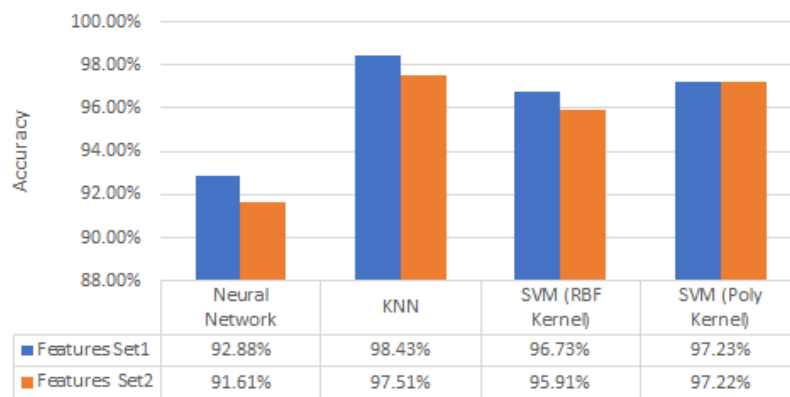


Fig. 21. Static dentist signs features comparison

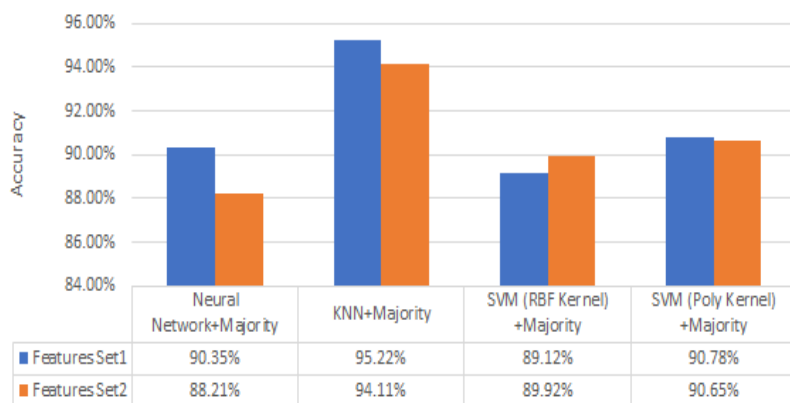


Fig. 22. Dynamic dentist signs features comparison

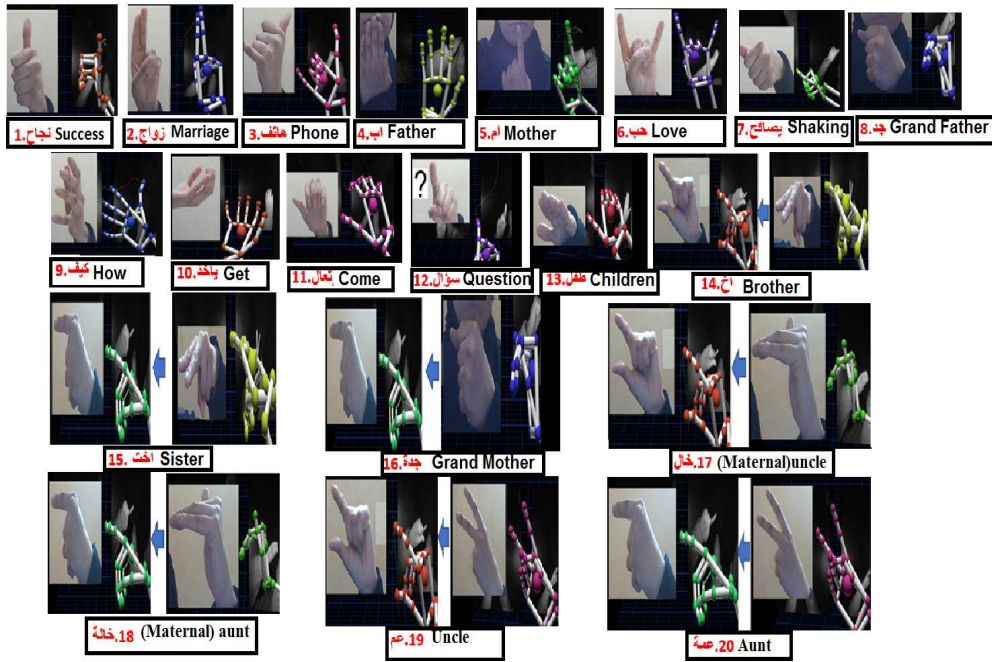


Fig. 23. Common verbs and nouns signs using one hand

Table 2. "DTW" results for dynamic dentist signs

Palm features set	98.5678%
Bone features set	98.0243%

Table 3. English meaning for common verbs and nouns signs with one hand

Arabic Sign	Meaning in English	Arabic Sign	Meaning in English
نجاح	Success	جد	Grand Father
زواج	Marriage	كيف	How
هاتف	Phone	ياخذ	Get
اب	Father	تعال	Come
ام	Mother	سؤال	Question
حب	Love	طفل	Kid
يصافح	Shaking	اخ	Brother
اخت	Sister	خالة	(Maternal) aunt
جدة	Grand Mother	عم	Uncle
خال	(Maternal)uncle	عمة	Aunt

Table 4. "DTW" results for dynamic gestures of common verbs and nouns with one hand

	Comparison
Palm features set	97.4512%
Bone features set	96.9712%

For Static Gestures, the training data set in classification method contains 50 samples for each sign from two different users and tested by 20 samples from another user, the comparisons between the used classifiers (KNN, SVM and ANN) in static common verbs and nouns with one hand are shown in Fig. 24.

For dynamic gestures, the results of the first method "Classification + Majority" and, the comparisons between the used classifiers (KNN, SVM and ANN) +Majority is are shown in Fig. 25.

For dynamic gestures, the results of the second method are represented in Table 4.

Common Verbs and Nouns with Two Hands

This experiment was performed using two hands and also contains static gestures such as (صديق(Friend), جار(Neighbour), منزل(Home), اجتماع(Meeting), متاسف(Sorry)) also there are dynamic gestures such as (طريق(Road), يعطف(CareFor), عائلة(Family), يعلق(Hang),

يخلط (Mix)). Some of them are static gestures and some are dynamic but both are performed by a single hand, they are shown in Fig. 26.

In this case, the number of features are 170 features for the Palm feature set and 140 features for the Bone feature set. The training set contains 50 samples for each sign from two different users and is tested by 20 samples from another user for both dynamic and static gestures. The equivalent English meaning are shown in Table 5.

For static gestures, the results of the first method "Classification + Majority", the comparisons between the used classifiers (KNN, SVM and ANN) + Majority is represented in Fig. 27.

For dynamic gestures, the results of the first method "Classification + Majority", the comparisons between the used classifiers (KNN, SVM and ANN) + Majority are represented in Figure 28 and the results of the second method "DTW" are represented in Table 6.

Table 5. English meaning for common verbs and nouns signs with two hands

Arabic sign	Meaning in English	Arabic sign	Meaning in English
صديق	Friend	يخلط	Mix
متأسف	Sorry	الطريق	Road
جار	Neighbor	يعطى	Hang
منزل	Home	يعطف	Care for
اجتماع	Meeting	عائلة	Family

Table 6. "DTW" results for dynamic common verbs and nouns with two hands

	Comparison
Palm features set	96.4128%
Palm features set	95.0143

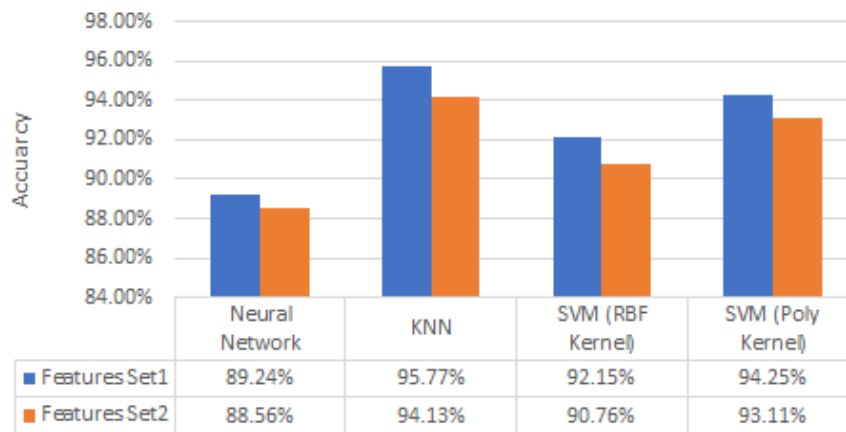


Fig. 24. Static common verbs and nouns with one hand features comparison

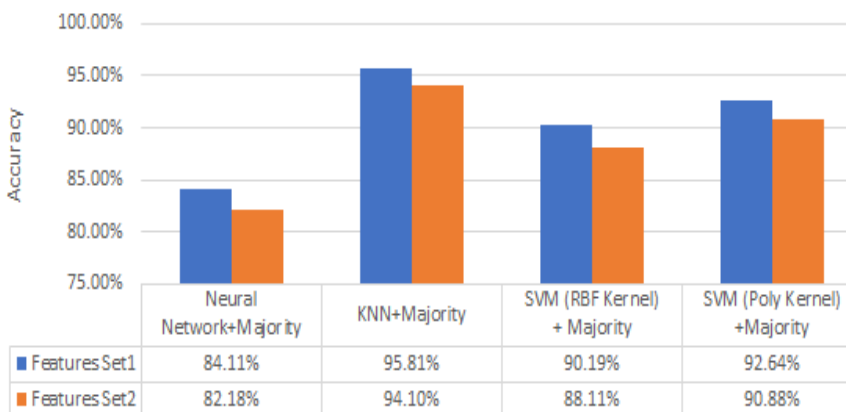


Fig. 25. Dynamic common verbs and nouns with one hand features comparison

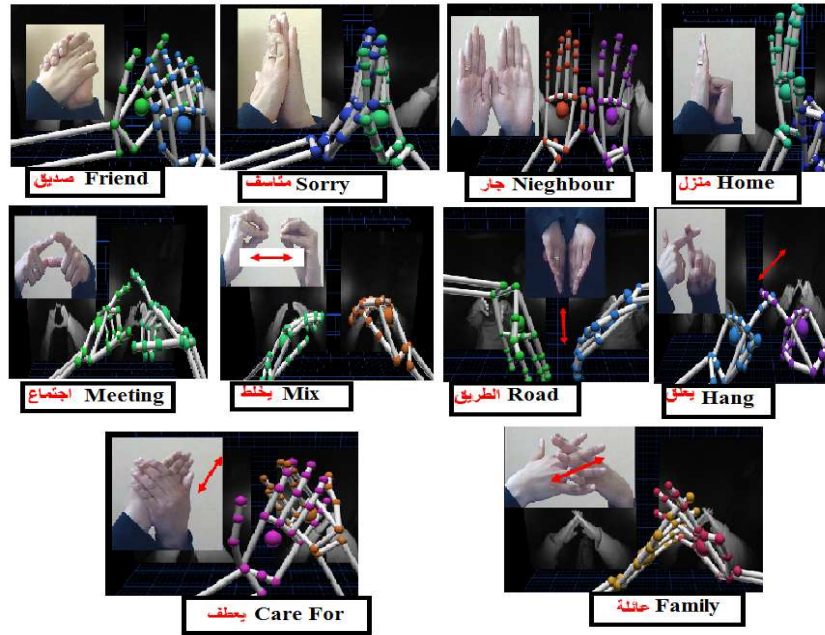


Fig. 26. Common Verbs and Nouns Signs using Two Hands

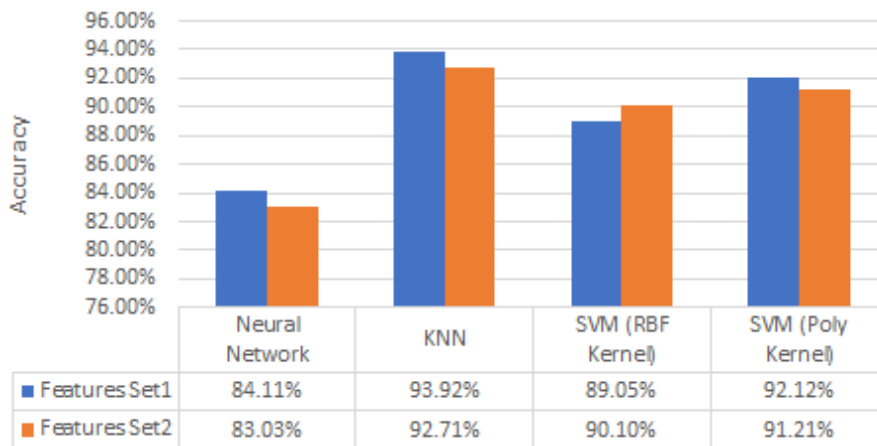


Fig. 27. Static common verbs and nouns with two hands features comparison

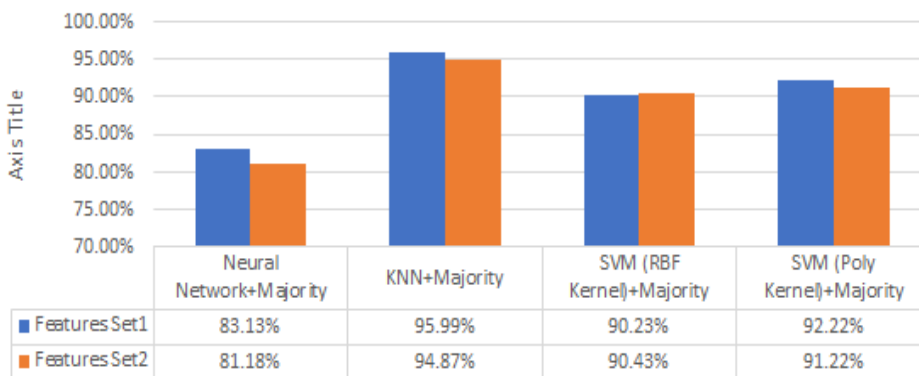


Fig. 28. Dynamic common verbs and nouns with two hands features comparison

Segmentation

We also worked on solving the problem of segmentation in real time in order to recognize the continuous sentences. After some experiments, we decided to use the palm speed during the motion as a segmenter. As it is shown in Fig. 29, there are difference in velocity when the transition between gestures or signs occurred, the segmenter will detect the changes in the velocity and perform the segmentation according to it. It is noticeable that the velocity is decreasing to less than 30 mm per sec.

when the gesture end then it increasing again over 100 mm per sec. when performing the next gesture. The segmenter keeps track the palm velocity and then segments the sequence of frames to groups each group contains a set of frames which represents a particular sign. Once the segmenter performs the segmentation, it runs the process of classification on each group individually to get the corresponding sign. The segmentation method was applied on sequences of numbers as in Fig. 29 which represents the palm speed over a sequence of numbers "523" represented by three signs "5" then "2" and "3".

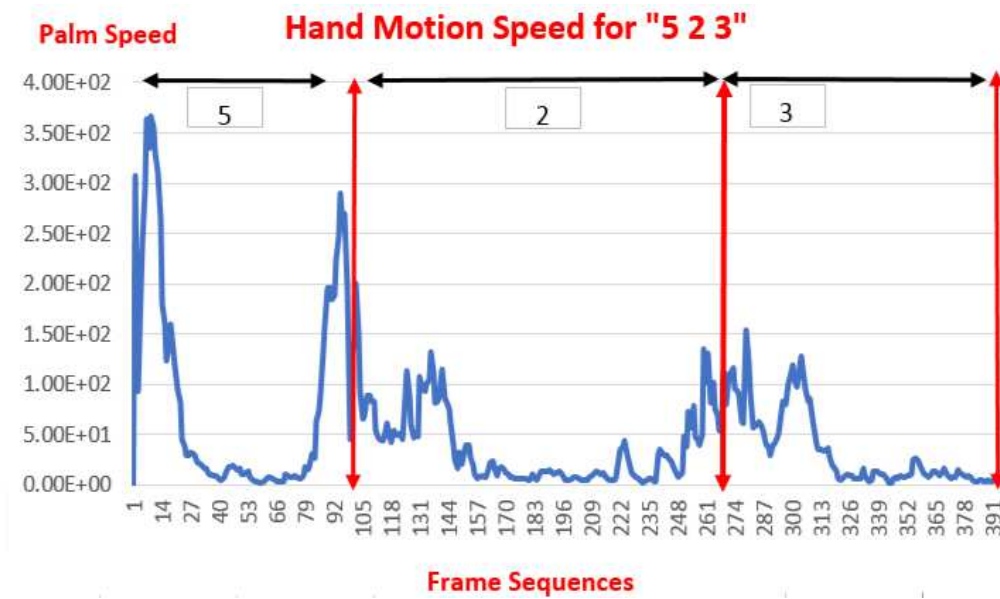


Fig. 29. Segmentation for numbers sequences

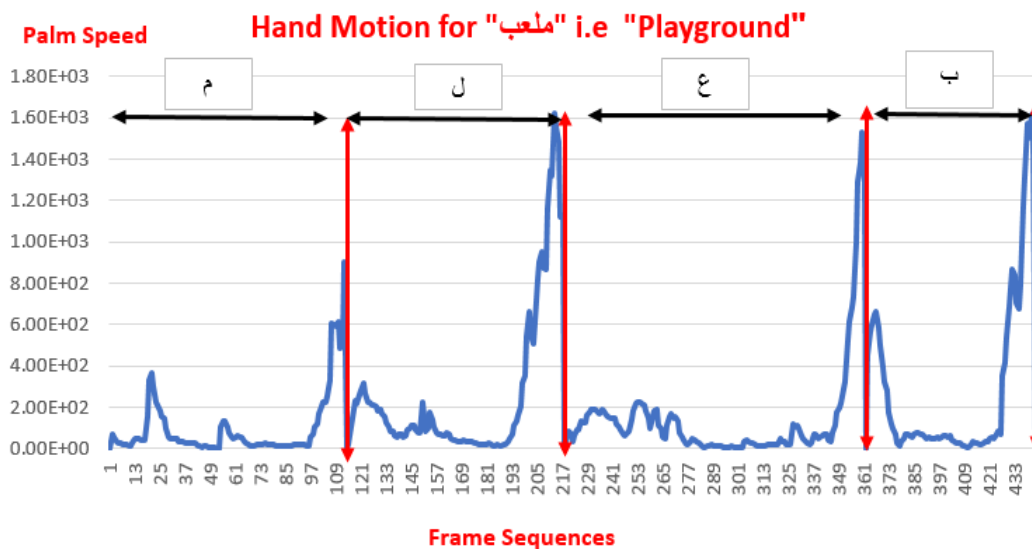


Fig. 30. Segmentation for finger spelling

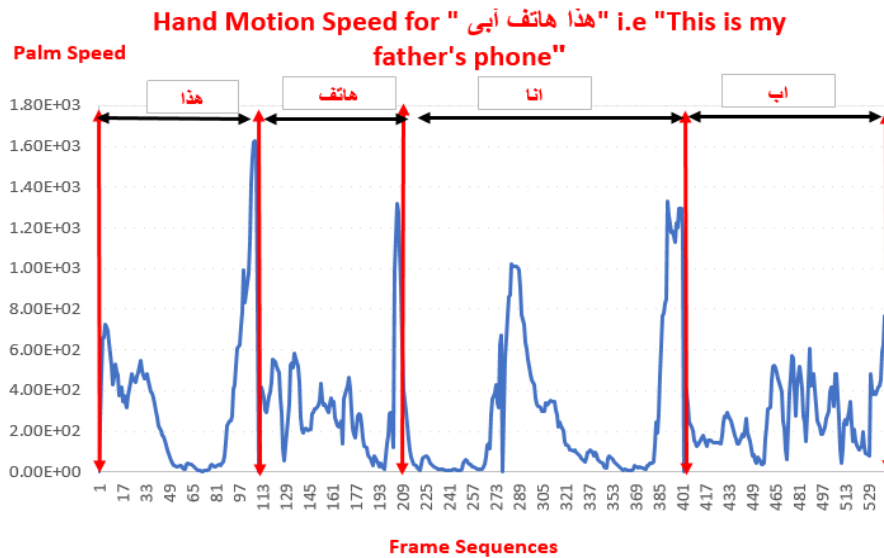


Fig. 31. Segmentation for sentence

Also, the segmentation method was applied on the finger spelling (i.e.) spelling the words with hand movement as in Fig. 30, which represents the palm speed over the word "ملاعب" i.e., "Playground" which is represented by four signs "م" then "ل" then "ع" and "ب".

Finally, we applied the segmentation method on the sentences which are consisting of several signs as Fig. 31 which represents the palm speed over the sentence "هذا هاتف ابي" i.e., "This is my Father's Phone" which is represented by four signs "هذا" then "هاتف" then "انا" and "اب".

We tested the segmentation on over 30 sentences with different lengths, the test concerned with the ability of this method to segment the sentences which were composed from several signs correctly and the results were over 95%.

Conclusion and Future Work

In this study, we developed a model for Arabic sign recognition using the leap motion controller (LMC). We applied our model on both static and dynamic gestures. Our experiments include the 28 Arabic Alphabets from أ to ع and Arabic Numbers from 0 to 10, eight common Arabic signs which are used at dentist, 20 common nouns and verbs used in the different aspects of life and finally 10 signs which are performed by two hands.

For static gestures, we used classification methods and compared the performance of three classifiers; Support Vector Machine (SVM) with poly kernel and RBF, K-Nearest Neighbor (KNN) and Artificial Neural Network (ANN) with a Multilayer Perceptron. Applying mentioned algorithms on two different features sets palm features set and bone features set.

For dynamic gestures, we suggested two methods: "Classification with Simple Majority" and "Dynamic Time Wrapping" which resulted in a good performance and accuracy of 98%. Also, the paper suggested a simple and effective solution to segment a series of continuous signs, which is the main problem in continuous recognition. This method depends on the motion speed, it works effectively in areal time and gave accuracy of 95%.

For future work, we plan to improve the accuracy of recognition by making additional features engineering and using deep learning with large samples of full sentences.

Acknowledgment

This research was supported by my supervisors Dr Ali Rashed and Dr Alaa Hamouda. We thank the reviewers for their so-called insights. We are also immensely grateful to their comments on an earlier version of the manuscript, although any errors are our own and should not tarnish the reputations of these esteemed persons.

Author's Contributions

Both the authors have equally contributed to this manuscript.

Ethics

This article is original and contains unpublished material. The corresponding author confirms that all of the other authors have read and approved the manuscript and no ethical issues involved.

References

- Al-Jarrah, O. and A. Halawani, 2001. Recognition of gestures in Arabic sign language using neuro-fuzzy systems. *Artificial Intell.*, 133: 117-138. DOI: 10.1016/S0004-3702(01)00141-2
- Aliyu, S., M. Mohandes, M. Deriche and S. Badran, 2016. Arabic sign language recognition using the Microsoft Kinect. Proceedings of the 13th International Multi-Conference on Systems, Signals and Devices, Mar. 21-24, IEEE Xplore Press, Leipzig, Germany, pp: 301-306. DOI: 10.1109/SSD.2016.7473753
- Aly, S., B. Osman, W. Aly and M. Saber, 2016. Arabic sign language fingerspelling recognition from depth and intensity images. Proceedings of 12th International Computer Engineering Conference, Dec. 28-29, IEEE Xplore Press, Cairo, Egypt, pp: 99-104. DOI: 10.1109/ICENCO.2016.7856452
- Assaleh, K. and M. Al-Rousan, 2005. Recognition of Arabic sign language alphabet using polynomial classifiers. *EURASIP J. Applied Signal Process.*, 2005: 2136-2145. DOI: 10.1155/ASP.2005.2136
- Benko, H., D. Wigdor, M. Haller, D. Lindbauer and A. Ion *et al.*, 2012. Understanding mid-air hand gestures: A study of human preferences in usage of gesture types for HCI. Technical Report.
- Biswas, K.K. and S.K. Basu, 2011. Gesture recognition using Microsoft kinect®. Proceedings of the 5th International Conference on Automation, Robotics and Applications, Dec. 6-8, IEEE Xplore Press, Wellington, New Zealand, pp: 100-103. DOI: 10.1109/ICARA.2011.6144864
- El-Bendary, N., H.M. Zawbaa, M.S. Daoud, K. Nakamatsu and K. Nakamatsu, 2010. ArSLAT: Arabic sign language alphabets translator. Proceedings of the International Conference on Computer Information Systems and Industrial Management Applications, Oct. 8-10, IEEE Xplore Press, Poland, pp: 590-595. DOI: 10.1109/CISIM.2010.5643519
- Elons, A.S., M. Ahmed, H. Shedid and M.F. Tolba, 2014. Arabic sign language recognition using leap motion sensor. Proceedings of the 9th International Conference on Computer Engineering and Systems, Dec. 22-23, IEEE Xplore Press, Cairo, Egypt, pp: 368-373. DOI: 10.1109/ICCES.2014.7030987
- Guesmi, F., T. Bouchrika, O. Jemai, M. Zaied and C.B. Amar, 2016. Arabic sign language recognition system based on wavelet networks. Proceedings of the IEEE International Conference on Systems, Man and Cybernetics, Oct. 9-12, IEEE Xplore Press, Budapest, Hungary, pp: 003561-003566. DOI: 10.1109/SMC.2016.7844785
- Hemayed, E.E. and A.S. Hassanien, 2010. Edge-based recognizer for Arabic sign language alphabet (ars2v-arabic sign to voice). Proceedings of the International Computer Engineering Conference, Dec. 27-28, IEEE Xplore Press, Egypt, pp: 121-127. DOI: 10.1109/ICENCO.2010.5720438
<http://blog.leapmotion.com/getting-started-leap-motion-sdk/>
<http://dartmouthbusinessjournal.com/2013/08/the-leap-motion-controller-and-touchless-technology>
<http://phys.org/news/2013-08-motion-ready-prime.html>
<http://www.who.int/mediacentre/factsheets/fs300/en/>
<https://developer-archive.leapmotion.com/getting-started/javascript/developer-guide>
https://en.wikipedia.org/wiki/Dynamic_time_warping
<https://www.leapmotion.com>
- Maraqqa, M. and R. Abu-Zaiter, 2008. Recognition of Arabic Sign Language (ArSL) using recurrent neural networks. First International Conference on the Applications of Digital Information and Web Technologies, IEEE Xplore Press, Aug. 4-6, IEEE Xplore Press, Ostrava, Czech Republic, pp: 478-481. DOI: 10.1109/ICADIWT.2008.4664396
- Mohandes, M. and S. Buraiky, 2007. Automation of the Arabic sign language recognition using the power glove. *AIML J.*, 7: 41-46.
- Mohandes, M., 2001. Arabic sign language recognition. Proceedings of the International Conference of Imaging Science, Systems and Technology, (SST' 01), Las Vegas, Nevada, USA, pp: 753-759.
- Mohandes, M., M. Deriche and J. Liu, 2014. Image-based and sensor-based approaches to Arabic sign language recognition. *IEEE Trans. Human-Machine Syst.*, 44: 551-557. DOI: 10.1109/THMS.2014.2318280
- Mohandes, M., M. Deriche, U. Johar and S. Ilyas, 2012. A signer-independent Arabic sign language recognition system using face detection, geometric features and a hidden Markov model. *Comput. Electr. Eng.*, 38: 422-433. DOI: 10.1016/j.compeleceng.2011.10.013
- Mohandes, M., S. Aliyu and M. Deriche, 2014. Arabic sign language recognition using the leap motion controller. IEEE 23rd International Symposium on Industrial Electronics, Jun. 1-4, IEEE Xplore Press, Istanbul, Turkey, pp: 960-965. DOI: 10.1109/ISIE.2014.6864742
- Mohandes, M.A., 2013. Recognition of two-handed Arabic signs using the cyberglove. *Arab. J. Sci. Eng.*, 38: 669-677. DOI: 10.1007/s13369-012-0378-z
- Quesada, L., G. López and L.A. Guerrero, 2015. Sign language recognition using leap motion. Proceedings of the International Conference on Ubiquitous Computing and Ambient Intelligence, (CAI' 15), Springer International Publishing, pp: 277-288. DOI: 10.1007/978-3-319-26401-1_26

Samir Elons, A., M. Aboul-Ela and M.F. Tolba, 2013. 3D object recognition technique using multiple 2D views for Arabic sign language. *J. Exp. Theoretical Artificial Intell.*, 25: 119-137.

DOI: 10.1080/0952813X.2012.680073

Shanableh, T. and K. Assaleh, 2011. User-independent recognition of Arabic sign language for facilitating communication with the deaf community. *Digit. Signal Process.*, 21: 535-542.

DOI: 10.1016/j.dsp.2011.01.015

Vijay, P.K., N.N. Suhas, C.S. Chandrashekhar and D.K. Dhananjay, 2012. Recent developments in sign language recognition: A review. *Int. J. Adv. Comput. Eng. Commun. Technol.*, 1: 21-26.