Smart hand gestures recognition using K-NN based algorithm for video annotation purposes

Malek Z.Alksasbeh[1], Ahmad H. Al-Omari[2], Bassam A. Y. Alqaralleh[3], Tamer Abukhalil[4], Anas Abukarki[5], Ibrahim Alkore Alshalabi[6], Amal Alkaseasbeh[7]
[1,3,4,5,6]Faculty of Information Technology, Al Hussein Bin Talal University, Ma'an, Jordan
[2]Department of Computer Science, Faculty of Science, Norther Border University, Arar, Saudi Arabia
[7]Ministry of Education and Higher Education, Al-Dawha, Qatar

Article Info

ABSTRACT

Sign languages are the most basic and natural form of languages which were used even before the evolution of spoken languages. These sign languages were developed using various sign "gestures" that are made using hand palm. Such gestures are called "hand gestures". Hand gestures are being widely used as an international assistive communication method for deaf people and many life aspects such as sports, traffic control and religious acts. However, the meanings of hand gestures vary among different civilization cultures. Therefore, because of the importance of understanding the meanings of hand gestures, this study presents a procedure whichcan translate such gestures into an annotated explanation. The proposed system implements image and video processing which are recently conceived as one of the most important technologies. The system initially, analyzes a classroom video as an input, and then extracts the vocabulary of twenty gestures. Various methods have been applied sequentially, namely: motion detection, RGB to HSV conversion, and noise removing using labeling algorithms. The extraction of hand parameters is determined by a K-NN algorithm to eventually determine the hand gesture and, hence showing their meanings. To estimate the performance of the proposed method, an experiment using a hand gesture database is performed. The results showed that the suggested method has an average recognition rate of 97%.

Corresponding Author:

Malek Zakarya Alksasbeh
Faculty of Information Technology
Al Hussein Bin Talal University, Ma'an, Jordan
Email: malksabeh@ahu.edu.jo

1. INTRODUCTION

The modern evolution of intelligent and smart applications in image processing technologies has provided significant services to the users. Normally, several human gestures are expressed by the body, head, face, arm, finger, and hands movement. However, hand gesture is the most famous and most commonly used assistive communication method. In addition, hand gestures have proven to be a robust means of communication to provide information [1].

Hand gesture recognition (HGR) is a challenging problem due to some important factors such as the unpredictable illumination conditions, uncontrolled environments, the relatively small size of the palm and the fingers, complexities of different signs, finger occlusions, and the overall complexity of the visual recognition of the hand gestures [2]. Therefore, our proposed system emphases on the recognition of sign words, which is a communications method using hands. This system helps in reducing the communication

gap between the deaf people and their community via developing an HGR system which tries to annotate sign-words being captured on a video.

The availability of annotated images or videos helps to establish a critical prerequisite for the implementation of an intelligent knowledge management services designed specifically to accommodate realistic user needs [3]. Image annotation is a procedure that mobilizes the idea of sticky note which is printed on images for illustrational purposes. Similarly, video annotations are implemented on video frames instead of images to generate a note in a specific place to be used later in the recognition process [4].

In this study, K-NN machine learning algorithm is used as the core classification method. Also, a classroom is employed as a case study scene. Twenty single-hand gestures were adopted as portrayed in Figure 1. The proposed methodology begins by importing the video from a camera device and then dismantling the video to frames. Then, each frame is analyzed to extract hand gestures. Thus, when the system detects the hand gestures, it assigns a note to it. This note appears as an illustration of what is being explained by the teacher as he talks. Finally, the reassembled video; along with the illustrational comments will be displayed on an output screen.

Figure 1. The single-hand gestures in this study

HGR system represents an ingenious user-friendly computer-based approach that is more appealing and easier to use by the human beings. A wide area of applications may utilize gesture recognition. These applications include human machine interaction, immersive game technology, sign language etc., [5]. HGR system might be a real time or an offline system.Lee and Hong [6]designed a real-time HGR system which is based on difference image entropy obtained using a stereo camera where hand detection has been firstly conducted in a constrained environment. More specifically, a recognition system is implemented for the incoming hand images in real-time. In the detection part of their system, an implementation of a depth map employing SAD method based on right-left images obtained using a stereo camera is performed. This system senses the foreground object and executes hand detection. Then, the HGR system employs the difference of the image entropy from both the input image and the average image. Experiments using a hand gesture database (size 240) were performedto estimate the performance of the proposed system. The results showed that the suggested method had an average recognition rate of 85%.

Panwar [7] described a different type of real time HGR system. This system is based on detecting some significative shape-based features such as orientation, centroid, fingers status, thumb in case of raised or bent hand fingers and their competent location in image. The adopted approach is fully reliant on the shape parameters of the hand gesture. The video acquisition stage utilizes a web cam which works on 20 fps with intensity of 7 mega pixel. Some pre-processing steps are involved to remove the background noise. K-means

clustering was employed to allow only hand objects or the segmented significant cluster to be processed as a previous step to preparing and computing shape-based features. This unsophisticated shape-based approach can recognize around 45 various hand gestures using 5-bit binary string, resulted as the algorithm output. However, the proposed algorithm testing was performed using 450 images. The results show approximate recognition rate of 94%.

Hikawa and Kaida [8] discussed a hardware posture recognition system using a hybrid network which consists of Hebbian network and self-organizing map (SOM). Input posture images are used to extract feature vectors, then these vectors are mapped to a lower dimensional map of SOM neurons. The Hebbian network is a feed-forward neural network with a single-layer trained with a Hebbian learning algorithm. The recognition algorithm is robust to the alteration in hand sign locations. However, it is not impregnable to alternation or scaling. Therefore, a noise was added to the training data used by SOM-Hebb classifier to enhance the robustness of the recognition algorithmagainst alternation and scaling. In addition, neuron culling is suggested to enhance the performance. A field-programmable gate array (FPGA) is used to implement the whole system. Additionally, a novel video processing architecture is implemented as well. The system was developed to identify 24 American sign languages where all of them were hand signs. Simulations and experiments, both were used to verify the system feasibility. The experimental results showed recognition speed of 60 frames per second and accuracy rate of 97.1%. The researchers offered an innovative hardware implementation and used a very small circuit size as recommended for embedded applications.

The researchers in [9] and [10] used a different type of sensors as hardware aid to develop a HGR system. Liuet al. [9] made the first effort to combine data from inertial and vision depth sensors. These data can be used for recognizing various body movements. Combining data from both sources was performedusing a complementary technique. This approach leads to a more robust recognition results compared with the positions when each sensor was used individually on its own. Similarly, Ren et al. [10] have offered new opportunities for human-computer interaction (HCI) via using Kinect sensor, which is a recently developed depth sensor. Robust HGRis still an open challenging research problem despite the excessive improvement that has been made by leveraging Kinect sensor since hand is relatively small object compared to the entire human body, with more intricate articulations and more readily influenced by segmentation errors. Consequently, theirproposed system focuses on constructing a robust part-based HGRsystem using Kinect sensor. Also, they proposed a novel distance metric called finger-earth mover's distance (FEMD) to figure the contrast between the hand shapesgained from the Kinect sensor in order to handlethe noisy shapes. The experiments showed that the proposed system is precise (a 93.2% mean accuracy on an opposition 10-gesture dataset), effective (average 0.0750 s per frame), robust to hand articulations, orientation and scale changes or distortions. Also,the system can work in unrestricted environments (pothered backgrounds and lighting conditions).

Xie et al. [11] proposed a smart ring based on an accelerometer and a matching similarity based extendible and gestures recognition algorithm. The hand motions accelerations are collected by the 3-axis accelerometer that is incorporated in the ring which can pursue gestures in 2-D space. This approach categorizes the gestures into two types, the basic gesture which is simple and the complex gesture which is sequence combination of basic gestures. An automatic identification for a sequence of individual gestures is presented in their work by performing segmentation on the input images. To identify the basic gesture, the proposed approach calculated the average jerk in order to extract effective features. Then, the identified gesture is encoded using Johnson code. At last, the complex gesture is identified by matching the similarity between the acquired basic gesture sequences with the stored templates. A library which contains both basic and complex gestures was created, while the users can simply set and insert their own gestures with no need for pre-training. The studied model achieves a rate of 98.9% for basic gesture recognition and a rate of 97.2% for complex gesture recognition. Comparing the resulted rates with complete matching, the algorithm based on similarity matching ameliorate the complex gesture recognition rate by about 12%.So, these results have successfully legitimized the effectiveness and feasibility of the gesture decomposition and recognition algorithm based on similarity matching.

Xu et al. [12] used an accelerometers which have presented a three different gesture recognition models. These models could recognize seven different hand gestures, i.e. left, right, up, down, tick, cross, and circle, based on the input signals which come from three MEMS 3-axes accelerometers. These accelerometers reveal handaccelerations in three perpendicular directions. Then, these accelerationsare transmitted via Bluetooth wireless protocol to a PC. An automatic gesture segmentation algorithm was improved to identify a sequence of individual gestures. Also, a basic feature was extracted based on sign sequence of gesture acceleration in order to compress data and to decrease the influence of dissimilarities resulted from gestures which made by different users. This approach reduced hundreds of single gesture data values to a gesture code of 8 numbers. Finally, the gesture recognition is performedby comparing the gesture

code against the stored templates. The findings of performing 72 experimentswhich contained a sequence of hand gestures (628 gestures in total) show that the superior model discussed in their workachieved an overall recognition accuracy of 95.6%. Also, the correct recognition accuracy of each gesture ranges from 91% to 100%. Concluding that sign sequence and template matching recognition algorithm can be employed for non-particular users' HGR without the need for the time-consuming process of training user prior to gesture recognition.

Another technique is given by Mazumdar et al. [13] presented a comparative study between two HGR methods, a gloved hand tracking and segmentation algorithm and a free hand tracking algorithm. The gloved hand tracking algorithm is unlike the free hand tracking in that it is not dependent on any background and lighting conditions. This robust and efficient gloved hand tracking system perfectly handled many problems such as skin color detection, the complexity added from the presence of many people in front of the camera, intricate background elimination and variable lighting condition. Presented noise in the segmented image by means of dynamic background can be eliminated with the aid of this adaptive technique which was found to be efficient for the application conceived.

Also, othervarious approaches with no hardware equipment have been introduced for the HGR problem. A system based on KL Transforms was proposed in [14] to identify different hand gestures. Five steps were adopted: filtering the skin, cropping the palm, detecting the edges, extracting the features, and classification. Primarily the hand was detected by filtering the skin, and cropping the palm was executed to extract out only the palm portion. In order to extract the outline images of palm, the extracted image was processed by performing the Canny Edge Detection technique. After palm extraction, the hand features were extracted employing K-L Transform technique, eventually the input gesture was recognized using proper classifier. The system was tested for 10 different hand gestures, and the obtained recognizing rate was 96%. In [15], the researchers measured the distance from the hand edges to the centroid of the hand, along a number of radials spaced equally around a circle. This informed that the general "shape" of the gesture.

Moreover, another seven different algorithms for hand feature extraction, were discussed in [16]. These algorithms aimed to recognize and furthermore classify static hand gestures. Gesture vocabulary was defined, with 10 gestures, and videos from 20 persons performing the gestures for hand feature extraction were recorded. The study aims to test the robustness of all algorithmsindividually. Each algorithm is applied to scale, translation then rotation invariance. The analysis of both data andobtained results indicated further pre-processing on the video frames is a must in order to reduce the number of variant feature values acquired for the same hand posture. The depth video images acquired with the Kinect had low resolution with some noise, so it was inferred that some impreciseness on data recordings results from those difficulties, leading to more complicated class learning. It was found that the radial signature and the centroid distance were the best shape descriptors discussed in this paper in terms of robustness and computation complexity.

Furthermore, two-color spaces into HS-CbCr format for skin detection and averaging background are combined to solve the background problem that was introduced in [17]. The experimental results show that the proposed combined method can recognize hand gestures in good lighting conditions.

In addition, computer vision approaches were used for HGR from a video stream as proposed in [18]. The proposed framework utilized both background subtraction and skin color models for segmenting each hand-palm region inthe video stream. Consequently, a fuzzy directional encoding strategy is employed on the polygonal approximation of a gesture to form features which are used to train an ANN classifier for developing a gesture prediction model.

Recently, an improved man-machine interaction by single person HGR system using Microsoft's Kinect sensor was presented in [19]. This system used skeletal images gained from the sensor. It can extract six features from 20 body joints' coordinates in 3D space. Gesture classification was performed using artificial neural network with back propagation learning algorithm. However, our proposed system performs four new K-NN parameters of machine learning pattern recognition field to get precise HGR for the twenty given gestures vocabulary which are used in this research. These parameters are discussed in detail in the next section.

2.   RESEARCH METHOD AND PROPOSED SYSTEM

There have been abundant researches in this domain. Several methodologies were suggested such as principle component analysis (PCA) method [20], subtraction method [21], K-Nearest Neighbor (KNN) machine learning method [22] and other HGR methods illustrated in [5, 23]. In the context of this research, KNN machine learning algorithm is preferred for many reasons such as simple implementation, ease of use, low calculation time, distinguished performance, and recognize gestures accurately [24-26]. Therefore, KNN machine learning algorithm is used as the core system in this study. A mock lecture in a classroom was recorded. The filmed video is imported and then divided into length, and width agreed frames, at this stage it

becomes easy to handle the video as a frame buffer, each of which is considered as an image that facilitate the analysis process. Analyzing and examining each frame is needed to extract the hand gesture (HG). If a HG exists in the frame, then the system tried to perform primary editing by adding a note which interprets the embedded gestures. The defragmented frames are reassembled to be displayed after the annotation process completed. The basic block diagram of HGR system is shown in Figure 2. To clarify the system in a more detailed format the following flow chart is established.

As portrayed in Figure 3 and Figure 4, the HGR system consists of the following modules: pre-processing, features extraction and post processing. MATLAB software has been used to develop and execute the following HGR system modules as an image processing tool.
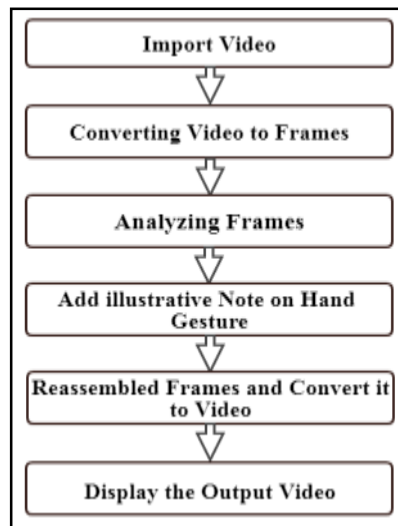

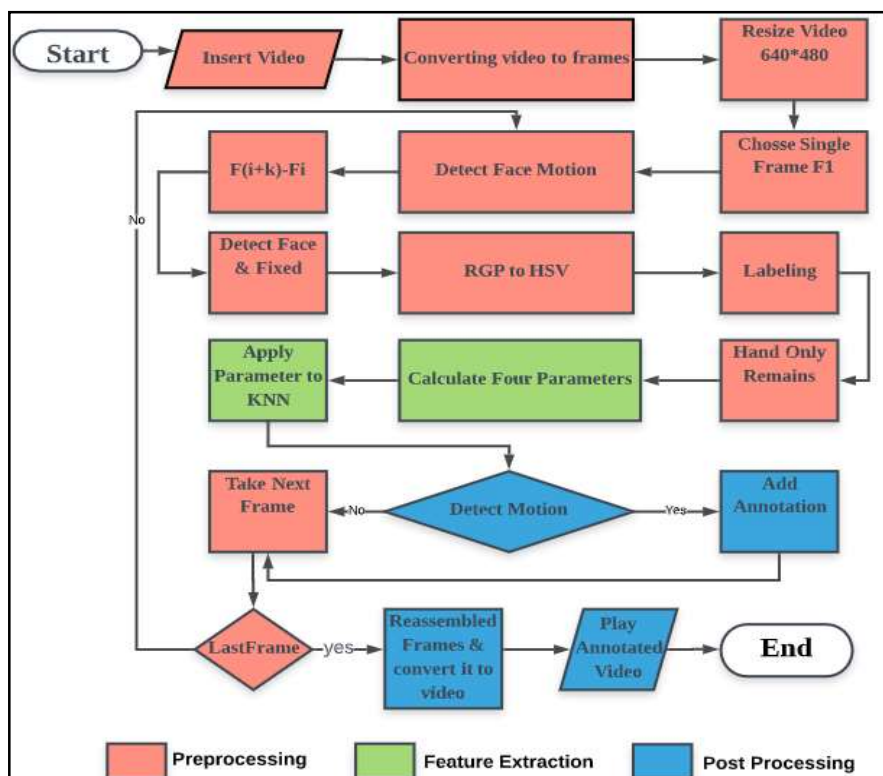Figure 2. Block diagram of the system
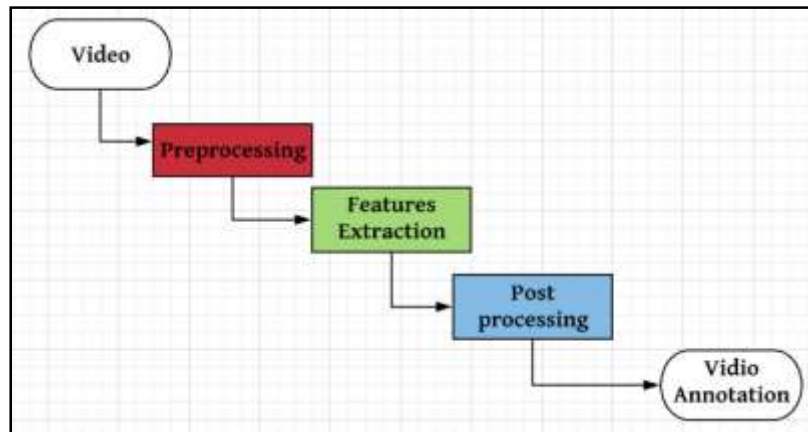

Figure 3. System flow chart

Figure 4. Simplified structure showing the main computational modules for HGR system

2.1. Preprocessing

This module allows the user to import/insert video in order to be processed. Basically, the preprocessing module consists of the following stages:

2.1.1. Converting video to frames

In order to simplify the processes applied to the video, it's important to fragment video into frames, since dealing with the video as frames is much easier to be handled.

2.1.2. Removing background and non-movable pixels

In this step, the system removes the background noise in order to eliminate any colour that could be similar to skin colour which may perturb the hand detection process. Thus, our system applies the subtraction method which subtracts pixel from input image to pixel in another image that is located in a number of frames far from the original frame, depending on the speed of teacher's motion, By doing so, our system neglects any fixed pixels, and keeps the movable ones including the hand.

2.1.3. Deleting the face

In this stage, the face detection function is applied in order to remove the face to eliminate any confusion might happen between hand and face.

2.1.4. Skin detection

The skin color detection is one of the major tasks in the HGR system. Skin color detection and decision principles are handled by the system. The system must take into consideration the discrimination between skin part and non-skin part pixels. This is achieved by introducing metric measurements, which measures the distance of the pixel color. This metric type is known as skin modeling. However, extracting movable pixels is not enough to recognize hand palm. Therefore, the system converts the resulted frame to the HSV color domain in order to detect skin color. There are plenty of models employed for skin detection such as RGB (Red, Green, and Blue), YCbCr (Luminance Chrominance) and HSV (Hue, Saturation Value). However, our system adopted HSV model which offer better results for segmentation applications than other methods [27, 28].

2.1.5. Removing noise

Image may have many skin surfaces seen by the camera. Since only one hand is required, the system performs an image *labeling* criterion in order to recognize the hand among all the skin regions. Label is essentially an integer values assigned to all skin regions pixels. When object had label, then assign the current pixel with the same label, if not then use a new label with a new integer value. After counting all labeled regions (divided image), we sort all of them in ascending order, and we take the area that have the maximum label value which we interested in because we assume that hand region is the biggest part in the image. For the sake of separating hand region, we create a new image which contain the value 1 in the positions where the maximum label occurs, and we set the other values to zero. Figure 5 shows the output of each step, from the time the video is processed until the hand is being extracted.
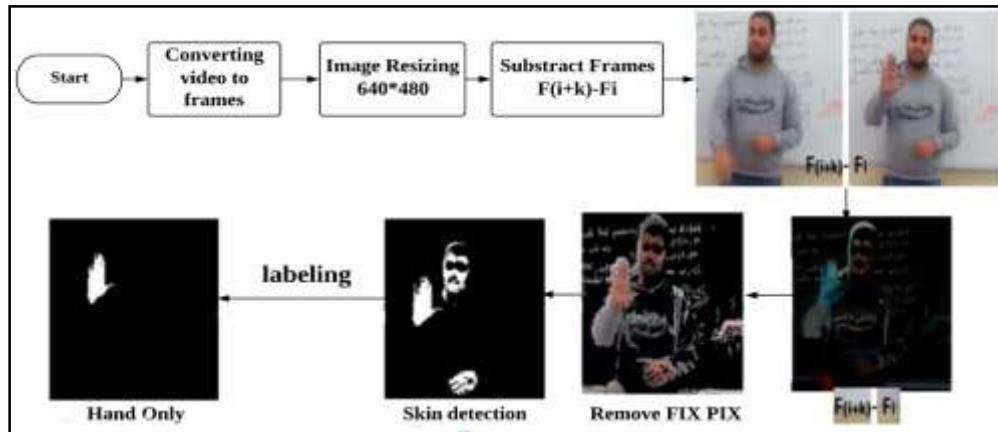
Figure 5. The sequenced output of each stage in preprocessing phase

2.2. Feature extractions

Employing KNN need entry parameters in order to extract the features and identify the hand region. The proposed system adapts four parameters for this purpose.

a) Lines ratio: the ratio between the vertical and horizontal lines; If (x_c, y_c) is the center point in the white hand shape,

$$\text{Vertical line} = x_c \tag{1}$$

$$\text{Horizontal line} = y_c \tag{2}$$

$$\text{Lines ratio} = \frac{Horizontal\ line\ length}{Vertical\ line\ length} \tag{3}$$

This ratio reflects hand propagation from right to left and up to down.

b) Rectangle ratio: for the rectangle circumstance that covers the hand shape and has the median point (x_c, y_c), suppose that summation S of hand pixels that intersects the rectangle is computed, then rectangle ratio can be calculated as follows:

$$\text{Rectangle ratio} = \frac{s}{Rectangle\ circumstance} \tag{4}$$

The mathematical ratio figures out the external boundary for the hand gesture. The previous two parameter ratios are illustrated in Figure 6.
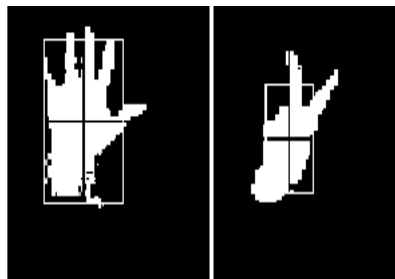

Figure 6. The ratio parameters for the supposed rectangle and lines

c) The second order column moment, where $\bar{c}$ is the rows' mean

$$\mu_{cc} = \frac{1}{Hand\ area} \sum_{(r,c)\epsilon R} (c - \bar{c})^2 \tag{5}$$

This parameter gives information about the horizontal distribution around the rows' mean within the whole hand area.

d) The standard deviation: is a measure of the data set dispersal from its mean. The more diffusion apart the data, the higher the deviation. Standard deviation is represented as the positive square root of variance.

If we have vector X which contains the number of hand pixels at each row for the hand gesture, where n is the last row number that is less than row's mean, then:

$$s = \sqrt{\frac{\sum_{i=1}^{n}(xi-row'smean)^2}{n-1}} \tag{6}$$

As shown in Figure 7, the standard deviation reflects the dispersal of hand density from its mean. The more diffusion scattering the data, the higher the deviation.

Figure 7. The standard deviation parameter effect

2.3. Post processing

At this point the final touch to the design is added. Post processing is the annotation that describes the hand gesture. Post processing allows the user to add his own annotation by building a graphical user interface (GUI), giving the user a choice whether or not to add any suggested annotation. Figure 8 shows the annotated frame as being recommended by the user.

Figure 8. Annotated frame

3. OBSERVE AND EVALUATE

The performance of the whole system is discussed in this section based on processing time, accuracy, sensitivity, and specificity.

3.1. Processing time

Time evaluation in image processing and recognition phases is very essential to estimate the performance. Processing time shows leaning in the technique which has been used to recognized hand gestures. There are few factors that discriminate the results such as image quality, image size (648x480 and 1920x1080) and the parameters of recognition technique or algorithm. In training, testing, feature extraction

and recognition of that particular image, processing time for our system is almost negligible since it achieves 2.1 ms/frame processing speed.

3.2. Effect of training pattern

This experimental assumption was made to reduce misclassification as the number of training patterns increase. There are many images of different people with skin color in the training database to enable KNN to learn in order to classify with different gesture positions. Mounting training pattern gives more effective and influential results.

In this study, confusion matrix is used to calculate the system accuracy, sensitivity, and specificity. A confusion matrix [29] contains numeric data about actual and result classification that is done by the recognition machine learning algorithm. The following Table 1 shows the form of confusion matrix. Where TP is the correct predictions of the positive instance, FB is the incorrect predictions of the negative instance, FN is the correct predictions of the negative instance and TN is the incorrect predictions of the negative instance.

Table 1. The confusion matrix format

		Predicted	
		Positive	Negative
Actual	Positive	TP	FP
	Negative	FN	TN

In this study, 120 Hand Gestured images for each class were used in these experiments. Then, accuracy, sensitivity, and specificity for twenty classes were calculated as portrayed in Table 2 and Figure 9 depending on the following formulas.

Accuracy (7) is how close a measured value is to the actual (true) value.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{7}$$

Sensitivity (8) which called the true positive rate measures the proportion of actual positives and is complementary to the false negative rate.

$$Sinsitivity = \frac{TP}{TP+FN} \tag{8}$$

Specificity (9) is the true negative rate reflects the proportion of negatives and is complementary to the false positive rate.
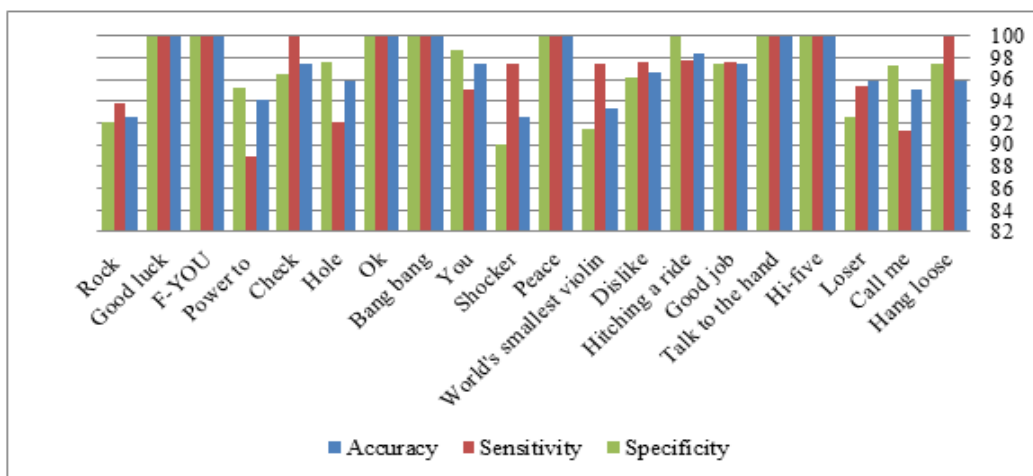
$$Specificity = \frac{TN}{TN+FP} \tag{9}$$

Figure 9. The accuracy, sensitivity, and specificity performance results

Table 2. The accuracy, sensitivity, and specificity results

Classes	Accuracy	Sensitivity	Specificity
Hang loose	95.83	100	97.43
Call me	95.00	91.30	97.29
Loser	95.83	95.34	96.10
Hi-five	100	100	100
Talk to the hand	100	100	100
Good job	97.50	97.56	97.46
Hitching a ride	98.33	95.45	100
Dislike	96.66	97.56	96.20
World's smallest violin	93.33	97.36	91.46
Peace	100	100	100
Shocker	92.50	97.50	90.00
You	97.50	95.00	98.75
Bang bang	100	100	100
Ok	100	100	100
Hole	95.83	92.10	97.56
Check	97.50	100	96.42
Power to	94.16	88.88	95.29
F-YOU	100	100	100
Good luck	100	100	100
Rock	92.50	93.75	92.04
Average	97.12	97.20	97.12

4. CONCLUSION

In this work, a description of a structured gesture annotation system is developed. the proposed system consists of algorithms implemented on single-person hand gesture recognition, As required in any research, we have performed an extensive research looking for various approaches for developing HGR methods. Four parameters were extracted to be the key-entries for K-NN classifier. The proposed system works in off-line mode. The system can detect single-handed individuals in the camera's field of view accurately, then it recognizes the gesture –when they appear successfully. GUI for video annotation is also presented allowing the user to add his own description on the video. The system performance is tested on 120 hand gestures for each gesture. The experiment results showed that the suggested method had an average recognition rate of 97%. As illustrated in this research the suggested system achieves excellent results and hits an acceptable accuracy, but for more promotions, there is some small issues needed to be solved such as, video resolution affects the processing time. Also, the proposed system works with single person HGR only, another challenge was the position of hand when doing the gesture, if it is in front of the face then the system will miss classify it. HGR is a wide science field which shows continuous upgrading systems and inventions, taking this critical fact into account and due to the important application fields, there are some suggestions that can be implemented in future work, such as processing an ONLINE video, detecting multi-person hand gesture, dealing with more hand and face gestures vocabulary, building big dataset of hand gestures for more efficiency and for Robotics and computer vision.

REFERENCES

[1]   M. A. Rahim, M. R. Islam, and J. Shin, "Non-Touch Sign Word Recognition Based on Dynamic Hand Gesture Using Hybrid Segmentation and CNN Feature Fusion," *Applied Sciences,* vol. 9, no. 18, pp. 3790, 2019.
[2]   H. Badi, "Recent methods in vision-based hand gesture recognition," *International Journal of Data Science and Analytics,* vol. 1, no. 2, pp. 77-87, 2016.
[3]   Q. De Smedt, "Dynamic hand gesture recognition-From traditional handcrafted to recent deep learning approaches," Sciences andTechnologies, Université Lille 1, 2017.
[4]   S. Feng, R. Manmatha, and V. Lavrenko, "Multiple bernoulli relevance models for image and video annotation." *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR),* 2004.
[5]   M. Panwar, and P. S. Mehra, "Hand gesture recognition for human computer interaction," in *2011 International Conference on Image Information Processing*, Shimla, pp. 1-7, 2011.
[6]   D. Lee, and K. Hong, "A Hand gesture recognition system based on difference image entropy," in *2010 6th International Conference on Advanced Information Management and Service (IMS)*, Seoul,South Korea, pp. 410-413, 2010.
[7]   M. Panwar, "Hand gesture recognition based on shape parameters," in *2012 International Conference on Computing, Communication and Applications*, Dindigul, Tamilnadu, 2012, pp. 1-6.
[8]   H. Hikawa, and K. Kaida, "Novel FPGA implementation of hand sign recognition system with SOM–Hebb classifier," *IEEE Transactions on Circuits and Systems for Video Technology,* vol. 25, no. 1, pp. 153-166, 2014.
[9]   K. Liu, C. Chen, R. Jafari, and N. Kehtarnavaz, "Fusion of inertial and depth sensor data for robust hand gesture recognition," *IEEE Sensors Journal,* vol. 14, no. 6, pp. 1898-1903, 2014.

[10] Z. Ren, J. Yuan, J. Meng, and Z. Zhang, "Robust part-based hand gesture recognition using kinect sensor," *IEEE transactions on multimedia,* vol. 15, no. 5, pp. 1110-1120, 2013.

[11] R. Xie, X. Sun, X. Xia, and J. Cao, "Similarity matching-based extensible hand gesture recognition," *IEEE sensors journal,* vol. 15, no. 6, pp. 3475-3483, 2015.

[12] R. Xu, S. Zhou, and W. J. Li, "MEMS accelerometer based nonspecific-user hand gesture recognition," *IEEE Sensors Journal,* vol. 12, no. 5, pp. 1166-1173, 2011.

[13] D. Mazumdar, A. K. Talukdar, and K. K. Sarma, "Gloved and free hand tracking based hand gesture recognition," in 1st International Conference on Emerging Trends and Applications in Computer Science, Shillong, pp. 197-202, 2013.

[14] J. Singha, and K. Das, "Hand gesture recognition based on Karhunen-Loeve transform," in *Mobile and Embedded Technology International Conference India*, pp. 365-371, 2013.

[15] W. Choudhari, P. Mishra, R. Rajankar, and M. Sawarkar, "Hand gesture recognition using radial length metric," *International Journal of Science and Research,* vol. 52, no. 1, pp. 865-869, 2014.

[16] P. Trigueiros, A. F. Ribeiro, and L. P. Reis, "A Comparative Study of different image features for hand gesture machine learning," in *7th Iberian Conference on Information Systems and Technologies*, Madrid, Spain, pp. 41-46, 2013.

[17] R. F. Rahmat, t. Chairunnisa, d. Gunawan, m. F. Pasha, and r. Budiarto, "Hand gestures recognition with improved skin color segmentation in human-computer interaction applications," *Journal of Theoretical and Applied Information Technology,* vol. 97, no. 3, pp. 727-739, 2019.

[18] S. Saha, S. Das, S. Debnath, and S. Banik, "A Hand Gesture Recognition Model Using Fuzzy Directional Encoding of Polygonal Approximation," *Emerging Technology in Modelling and Graphics*, pp. 217-229: Springer, 2020.

[19] B. Ganguly, P. Vishwakarma, and S. Biswas, "Kinect Sensor Based Single Person Hand Gesture Recognition for Man–Machine Interaction," *Computational Advancement in Communication Circuits and Systems*, pp. 139-144: Springer, 2020.

[20] D.-Y. Huang, W.-C. Hu, and S.-H. Chang, "Vision-based hand gesture recognition using PCA+ Gabor filters and SVM," in *2009 fifth international conference on intelligent information hiding and multimedia signal processing*, Kyoto, pp. 1-4, 2009.

[21] A. Licsár, and T. Szirányi, "Dynamic training of hand gesture recognition system," in *Proceedings of the 17th International Conference on Pattern Recognition*, 2004. ICPR 2004., Cambridge, pp. 971-974, 2004.

[22] Y. Liu, X. Wang, and K. Yan, "Hand gesture recognition based on concentric circular scan lines and weighted K-nearest neighbor algorithm," *Multimedia Tools and Applications,* vol. 77, no. 1, pp. 209-223, 2018.

[23] M. M. Hasan, and P. K. Mishra, "Hand gesture modeling and recognition using geometric features: a review," *Canadian journal on image processing and computer vision,* vol. 3, no. 1, pp. 12-26, 2012.

[24] S. Zhang, X. Li, M. Zong, X. Zhu, and D. Cheng, "Learning k for knn classification," *ACM Transactions on Intelligent Systems and Technology (TIST),* vol. 8, no. 3, pp. 1-19, 2017.

[25] G. Guo, H. Wang, D. Bell, Y. Bi, and K. Greer, "KNN model-based approach in classification." *Springer*, pp. 986-996, 2003.

[26] M. M. Otoom, M. Jemmali, Y. Qawqzeh, K. N. SA, and F. Al Fay, "Comparative Analysis of Different Machine Learning Models for Estimating the Population Growth Rate in Data-Limited Area," *IJCSNS,* vol. 19, no. 12, p. 96, 2019.

[27] D. J. Bora, A. K. Gupta, and F. A. Khan, "Comparing the performance of L* A* B* and HSV color spaces with respect to color image segmentation," *International Journal of Emerging Technology and Advanced Engineering,* vol. 5, no. 2, pp. 192-203, 2015.

[28] S. Toure, O. Diop, K. Kpalma, and A. S. Maiga, "Best-performing color space for land-sea segmentation," in *2018 41st International Conference on Telecommunications and Signal Processing (TSP)*, Athens, Greece, 2018, pp. 1-5.

[29] F. Provost, and R. Kohavi, "Glossary of terms," *Journal of Machine Learning,* vol. 30, no. 2-3, pp. 271-274, 1998.