ISSN 2278-3091
Volume 9, No.5, September - October 2020
International Journal of Advanced Trends in Computer Science and Engineering
Available Online at http://www.warse.org/IJATCSE/static/pdf/file/ijatcse114952020.pdf
https://doi.org/10.30534/ijatcse/2020/114952020

Performance analysis and high recognition rate of automated hand gesture recognition though GMM and SVM-KNN classifiers

[1]Anuradha Patil, [2]Dr. Chandrashekhar M. Tavade
[1]Assistant Professor (ECE department) Godutai engineering college for women, kalaburagi, anuradha.keshul@gmail.com
[2]Principal at Basavakalyan engineering college, Basavakalyan, Bidar, cmttc1@gmail.com

ABSTRACT

The significant study about a hybrid framework design for hand gesture recognition of diminishing and disabilities in human body in daily life is increasing and it became a challenging task for identifications of gestures. There are various biomedical methods for identifications of hand gestures and organs in body, one of the best methods for identification gesture, neuromuscular abnormities is EMG. The framework has used a database containing RGB and Depth data, the study represents the review on Hand gesture recognition utilized in recognition of sign language. The hand gesture recognition is the prime research objective in HCI (human computer interaction) field, since of its significant application in sign-language recognition and computer games. Therefore, prime objective is to developing a hybrid framework which can provide the performance analysis for RGB and DEPTH data for hand gesture recognition system in high cost effective manner by considering of ASL. The proposed framework majorly comprises four phases 1) based on shape another then each hand unique feature 2) Training-phase which will classify the hand gestures with help of live video capture from Kinect camera using SVM which is linear and K-NN with weights, 3) Extraction of features using fingertip and web locations estimated and 4) Segmentation of gestures by using Gaussian Mixture Modeling (GMM). Overall, the algorithm achieved moderate accuracy and speed performance under most conditions. Additionally, it appears mostly insensitive to differences between individual hands, focusing more on the overall shape rather than unique features of each hand.

Key words : Hand gesture recognition, SVM, HMM, K-NN, Human computer interface, EGM.

1. INTRODUCTION

To create effective communication between computer and person, hand gesture system is employed. For transmitting useful data and for controlling a Robot these recognized hand gestures are utilized [1]. The interaction between computer and person is more accurate and is called "Human-Computer Interaction" which is also termed as "Man-Machine Interaction". But the machine has become trivial without the appropriate usage of human [2][3]. Performance and availability are the two main features of observed while designing a Human-Computer Interaction system. Performance is referred to as the facilities provided by the system to the user. System availability refers to how the system can control and perform specific tasks of the user [1] [4].

To interact between machines and person gestures are utilized and for interacting between person sign language is employed [5][6]. There are two types of gestures one is static (position / specified posture) another one is dynamic gestures (sequence of positions). Computational complexity required by the static gestures is less when compared to that of dynamic gestures. To obtain basic information regarding the gestures recognition system various techniques have been established [4] [6]. Some methods are utilized for external hardware-devices such as data-glove devices and color-markers which can easily extract the comprehensive-description of Gesture-features. To easily obtain the complete details of gesture-features some approaches have been employed in external hardware devices like data-glove devices and color-makers and basic features are extracted with the help of the appearance of the hand. These approaches are of less cost and simple when compared to other methods [7] [8].

The factors that influence “perceptual user interface (PUI)” are recognition of lip movement, tracking of eyes, voice, and gesture recognition of face [9]. The main objective of PUI is to enhance performance. Gestures are commonly used for input information in the personal-computing system [8] [9]. When the computers are physically damaged these recognized gestures are more useful and provides effective communication in the 3D-virtual environment or gambling [10]. The device that can enhance hand & body gestures comprises “Accelerometers” and “Gyroscopes” which can detect variation, orientation and increase the speed. Specific gestures can be interpreted with the help of a camera that is placed in the computing-device so that the software in the device can detect and identify the gestures [11]. Sign-language is considered as one of the useful devices for communication. For the people who are unable to hear but can talk and also for those who can hear and unable to talk and vice versa, these sign-languages are very helpful [5]. Sign-language includes various gestures, facial expressions, hand and body shapes, and movement. Deaf people show their feeling by making use of these types of gestures. Every facial expression, hand, and body movement provides a specific gist. Various sign-languages are used in different countries [6]. For instance, “American Sign Language” (ASL) is a common language employed in the USA, "Japanese sign language" (JSL) is used in Japan, "Indian sign language" (ISL) is used in India and French "French sign language" (FSL) is employed. Research on ISL has just started after the standardization of Indian sign language [6].

In Standard English language which is a constituent of 26 different alphabets which directly requires 26 hand gestures which are corresponding user hand shapes making the manual set each alphabet. The combination of the alphabets will result in the appropriate words and words leads to sentences and these resulted combinations of words and sentences will land in more than 6000 gestures in line with the ASL and this sign language is related with the various body area movements as shapes following the exterior appearance of the body shapes but without considering the appearance to provide the meaningfulness of the words and sentences which are supposed to be communicated. To obtain Hand Gesture information, several mechanical devices were employed in the starting stage of HG detection. "Data Glow" is considered as one of the highly accepted technique for Hand Gesture Recognition. The computational performance has been enhanced due to the development of computer technology. Non-wearable devices can satisfy user’s requirements with computer interface because of its significant features. Therefore, data gloves were substituted by other non-wearable devices [14]. The hand movement performance is reduced by the Data Glow process [15]. Due to scalability, robustness, and efficiency Hand Gesture Recognition system based on the computer vision process is gradually becoming popular.

The recognition stage plays an important role in the Hand Gesture Recognition system. This system requires an appropriate model of hand so that the human-computer interface (HCI) can accept the image. Hand kinematic structure is considered as a basis for the modeling approach to perform hand modeling [2][16][17]. Temporal modeling and spatial modeling are the two types of hand modeling. Temporal modeling deals with active or hand movement features in real-time whereas gesturing features an HCI environment. Over 2D and 3D space, the spatial domain can be employed [2] [3]. Based on the shape, movement, and deformable patterns, the 2D modeling of the hand is considered. Non-geographic and geographic models are the two types of shape-based hand modeling [2] [3]. Non-geographic models are utilized for feature extraction. This model employs shape-based features of hand such as shape, edges, Eigenvectors, etc. Even if there is a slight difference in the hand shape, the object can be easily changed in flexible/deformable models.

The 3D model of the hand is described and categorized as three types. They are geometrical models, volumetric models, and skeletal models. The geometric model is employed in hand animation and real-time applications. The simulation of the visual hand image is extensively performed on geometric surfaces. Geometric models consume more time as it requires more parameters. To shape the hand, only a few parameters are required by the skeleton models. Whereas more parameters are required by volumetric models to shape the hand and are more complex [14, 16]. The visual shapes such as ellipsoids and cylinders are substitutes for geometric shapes [17]. A basic scheme to combine data from inertial and vision framework depth sensors was analyzed by (Liu et al, 2014) for hand gesture recognition which employs the Hidden Markov Model (HMM) in its structure. Recognizing different movements from various parts of the body is the concept presented in this work. By combining the features of the vision depth sensor and inertial body sensor, invalid data produced from one sensor is counterbalanced by the other sensor. HMM illustrates about observation probability matrix and state transfer probability matrix. HMM also maps random processes [21].

Dynamic Time Warping (DTW) and Elastic Matching (EM) are the two approaches used for recognizing hand gestures. 93% is the Overall recognition rate achieved. Recognition rate when operated separately was found to be more. To create an interface between other gestures for recognition purpose this framework has been extended [24]. Micro-Doppler signatures are a technology presented by Kim et al, in 2016 which is projected by Doppler radar with the help of a deep convolutional neural network (DCNN). The main drawback of the optical recognition system of gestures is that it requires light rays every time whereas this issue is overcome by the Doppler system [22]. By employing micro-Doppler signatures ten hand signals can be analyzed such as to rotate, swipe, and push without any limit. Finger movement spectrograms were studied to examine the micro-Doppler signatures in a short-time fast Fourier transform. 90% of the data was used for training purposes and 10 % to isolate the spectrograms through DCNN. 85.6% was the efficiency of the technology when computed for ten gestures and the efficiency calculated was more i.e. 93.1% for seven gestures. These seven gestures which achieve high efficiency is implemented in real-time applications by using the proposed Doppler radar method. Based on the angle and distance from the radar, the Micro Doppler radar signatures vary [25].

Maqueda et al, in 2016 has introduced a technique for the extension of original volumetric spatiograms of local binary patterns (VS-LBP) in local binary patterns (LBP) descriptor which is used to enhance the efficiency and hand gestures are encoded that are related to temporal and spatial information. While obtaining feature extraction there exist many problems like different perspectives, occlusions, modifications, rotations, illuminations. Finger-Earth Movers Distance (FEMD) is a method implemented to match the templates. This analysis consists of a novel descriptor for hand-gesture recognition considered in sequences of videos. For real-time applications, feature descriptor is not feasible because of its large dimensionality. Therefore, temporal pyramid matching (TPM) of LBP is employed which can vary the spatial pyramid matching to a temporal domain. The idea of Local binary subpatterns is computed by minimizing the dimension of the descriptor. At last, to obtain hand pose representation, the final video descriptor combines local and global spatial information. The result obtained by the proposed system is more feasible and enhances the capability to recognize the dynamic hand gestures [26]. The four hand gestures the program was designed to recognize are listed in Fig.1 shows the possible combinations of hand gestures.

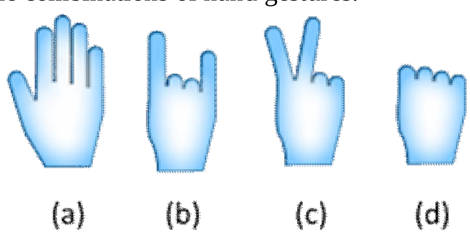

Figure 1: Four possible hand gestures through Kinect camera

2. PROBLEM STATEMENT

"Hand gesture recognition" is considered as one of the major drawbacks in the field of Human-computer interface because it is widely used in applications such as sign language recognition, computer games, virtual reality. Although, earlier works have developed a robust hand gesture recognition system developing real-time applications continues to be a challenging task. Current vision-based methods are significantly confined by the quality of the input image from the optical camera. The main causes for reducing the quality of the image are background clutters and variations in lighting. As a result, expected results are not produced by these systems for hand gesture recognition which is confined by the quality of input image obtained from optical cameras [18] [19] [20].

- Hand gesture recognition consists of two challenging issues: One of the main challenging tasks of hand gesture recognition is hand detection i.e. how actively one can detect the hand. The second issue is gesture recognition that is how accurately and effectively the hand gestures can be recognized. Fig.1.3 shows the architecture of the designed hand gesture recognition system. Usually, conventional approaches make use of color-makers for detecting hand. But in this proposed methodology, the handshapes are detected with the help of a "Kinect sensor camera" which is a unique camera consisting of a sensor that is used to obtain the depth map and color image and provides robustness to background clutters. The segmented handshapes are denoted as time-series curves
- Although, Kinect sensor has a resolution of only 640X480 large objects can be tracked by it. Consider a human body, to identify and segment small objects from an image with this resolution accurately is a difficult task i.e. in an image of the human body hand occupies a very small portion. Therefore, employing Kinect sensors remains a challenging issue for gesture recognition. Hence to overcome the differences between various hand shapes, a new shape distance metric termed as Finger-Earth Mover’s Distance (FEMD) is proposed in [20]. To match the hand shapes, FEMD metrics are designed which can overcome distortions and dissimilarities in hand.
- Handshape is denoted by signature in FEMD where each finger is considered as a cluster, and the distance variation between two handshakes is defined as the sum of the work required to move the earth piles and the consequence on the unequaled fingers. Two new finger detection algorithms are proposed to accurately detect the fingers. The first algorithm employs thresholding decomposition and the second algorithm is utilizing a near-convex shape decomposition scheme. At last, using a template matching the input hand is recognized.

3. PROPOSED METHODOLOGY

The study and design process of recognition of American-Sign-Language (ASL) through Kinect is specified in this proposed methodology which makes use of a bag of features (BOF). Communication by gestures is identified with the help of Kinect-sensor. The main drawback of the proposed strategy is Heterogeneous background i.e. the skin color corresponds to cloth color which possesses low image recognition and poor segmentation. Therefore, to overcome these issues some important steps are followed they are 1) create a database, 2) test BOW, 3) train the DataBase, 4) Gesture recognition and 5) final recognition. The complete details regarding the above steps are explained in the following section with the help of the flow-chart, algorithm, and resulting outcomes.

3.1 PROPOSED SYSTEM WORKFLOW

Since human-computer interaction is a prerequisite, the Hand signal process had become a challenging task in the early decades. In US states, American Standard Language is considered as one of the main sign languages for deaf peoples.

Basically, in US-countries ASL is generally studied as a second language. An ASL consists of several fundamental expressions such as face movement, hands, and upper body movements. In this proposed system, we study the ASL which deals with the various manual set of letters that make use of 26 hand figures. Recognizing the hand gestures of ASL is the primary goal of this research work. The gesture (ASL) symbols are recognized once Hand Direction, introduction, and Area which are considered as the manual parameters for gesture-based communication are utilized. In this proposed system, we make use of the RGB+D set of databases for hand gesture recognition system which consists of 26 alphabetical symbolic images i.e. from A to Z. In the proposed system, a Kinect camera is used to capture RGB-image data and Depth-image data. From the bag of features (BoF), the Kinect camera extracts some strong features of the image. To study the shape of the particular sign RGB-image and D-image are mapped. With the help of this data, hand alignment is carried out and hand movement is transferred to the cloud. Feature extraction is done by extracting the features from BoF. By using K-means clustering method features are extracted by aligned RGB-image and D-hand image. Robust features of the hand are obtained by using the SURF feature extraction method. After feature extraction process entire data samples (ASL) are trained and classification of images are done with the help of SVM classifier and these images can be considered for various experimental purpose. In the end, the images are recognized and validated.

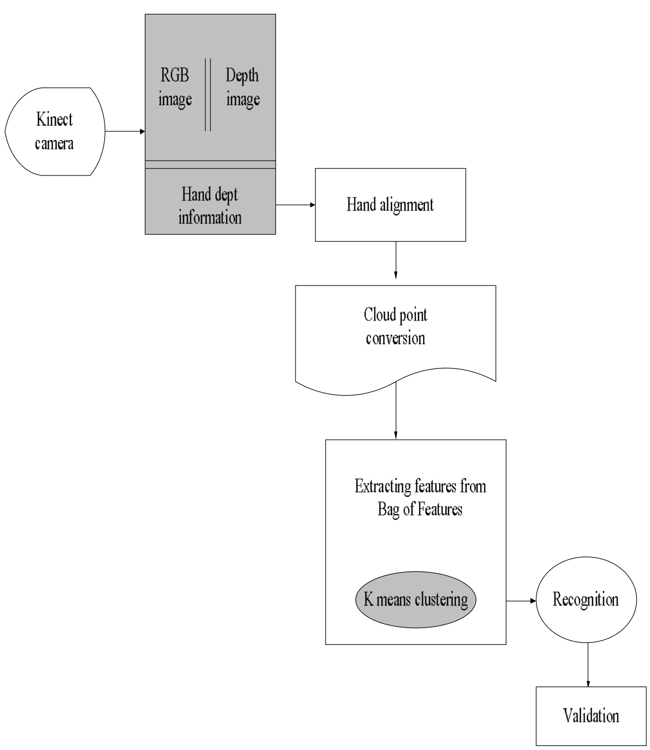

Figure 2: Proposed Flow Diagram of hand gesture recognition

The above-proposed execution process is initiated with the help of the database. This database includes 1250 RGB+D images along with 24 alphabetical symbols. The user-defined directory consists of American standard language alphabetical symbols that are present in the RGB+D image. For further process, the user is required to choose the database from the directory defined by the user. The following algorithm signifies the procedure for creating a database to select RGB and D-image. For creation of dataset, the following notation are used for calculation of depth and recognition rate

D-Sensor : for Depth calculation
Ix : Input image
D-image : Depth image
R_H : Right hand
Rx, Ry : Right x-axis and y-axis
Bw : Hand region

A. *Algorithm: CREATE DATABASE*

1. Start
2. Assign Fname = ‘xyz;
3. Define no, of image for each class
4. Delete webcam objects (if it’s already accessing)
5. Create an object of color and D-sensors.
6. Get srcDepth for depth device
7. Set mode (tracking to the skeleton)
8. Set frames for both devices →1
9. Continue set trigger for both devices → 200
10. Configure the camera for triggering for both sensors
11. Start video objects for both devices.
12. Set flag = 0
13. For (i = 5: -1: 1)
    a. Trigger both objects
    b. Obtain the acquired frames and metadata
    c. Obtain track ID
    d. Determined which ID is tracked
    e. Get x, y, and z coordinates
    f. Show image (“color image”)
    g. Show (Depth image)
    h. If ~(ix) empty
    i. Get D-image of R_H
    j. GetR_H coordinates (Rx and Ry)
    k. Plot (Rx, and Ry)
    l. Segment the Hand
    m. Consider Hand region (bw)
    n. End
    o. Pause (1)
14. End
15. Load idxx
16. Repeat step (13) for i = 1 to no_images
17. Save idxx, idxx

3.2 HAND ALIGNMENT

With the help of Point cloud library v.1.6.0, the segmented part of Hand is employed as a binary mask over D-image as soon as hand segmentation of RGB and D-image is performed. Background data is obtained by selecting the threshold which can remove background contextual objects. The system furnishes the amplitude value obtained along with the cloud which illustrates the potency of the active illumination signal considered by the object. In this step, the color image pixel is mapped to depth data using calibration parameters proposed in [58]. By employing calibration parameters, the depth data is mapped with image pixel which is proposed in [58]. Each pixel x-d and y-d is calculated through metric-3D by the depth camera. RGB-D 2D image is obtained by projecting a 3D point on the color image with the help of a built-in depth camera. By utilizing the distortion coefficient undistorted RGB-D image is predicted. The angle is non-negative when the hand is placed towards the right side and if the angle is negative, then the hand is positioned towards the left side. Therefore, to overcome rotation and translation problems the hand is aligned in the direction of the Y-axis, and also, we estimate the main diagonals and corresponding angles.

3.4 EXTRACTING FEATURES FROM BAG OF FEATURES (BOF)

The K means clustering method has to be considered for Feature Extraction. Local feature clusters can be easily obtained by this method, RGB+D-vector is generated with the help of specific data obtained and the number of histograms of cumulative-magnitudes is concatenated with this vector. The steps required to generate RGB+D-vector is explained by the following algorithm. The below algorithm shows the steps for Feature extraction from Bag of features with the help of K means clustering approach.
For calculation of Feature extraction by using BoF, the following notations are used

CPdept : Cloud point depth
RGB+D-vector : Depth vector
Si : sub-regions
Cp : Central point

B. *2. Algorithm: Feature-Extraction*

Input: CP_{dept}

Output: D-vector

1. Begin
2. The cloud point (CP_{dept}) is split into (n x n) sub-regions (S_i)
3. The central point (Cp) is evaluated for (S_i)
4. Generate the directional-vectors(d_{vector}) between (Cp) and (Pd) points
5. For each (d_{vector}) in (S_i), calculate its magnitude $|d_{vector}|$
6. Obtain the angles, x, y, and z
7. The cumulative magnitude orientation-histograms is calculated for each (S_i) and coordinates (X, Y, Z)
8. In each (S_i), three cumulative magnitude histograms (Hx, Hy, and Hz) is generated
9. Finally, Local feature D-vector is created and concatenating with cumulative histograms from each (S_i)
10. End

4. RECOGNITION

The device recognizes the input image as soon as feature extraction is carried out with the help of input images. These images make use of Bag of features from K means clustering. Permitting Non-linear classification is one of the major aspects of the SVM-classifier. With the help of the SVM-classifier, better results are obtained in predicting the gesture image. To study the performance of the designed system 40% of the data is taken into account for the training set and by comparing the proposed method with the current approach, the average accuracy of the approach is 97%. The execution of the proposed hand gesture recognition system is simulated on each image in the dataset having 103 images has been timed in chunks corresponding to the major stages of the algorithm: segmentation, feature extraction and classification. For the purposes of this test, image capture and plot creation time were ignored, in order to evaluate the algorithm itself, rather than the visualization code. The results of the speed test are shown in Table 1 for all three proposed techniques in terms of mean, standard deviation minimum and maximum and all in milliseconds.

Table.1 performance analysis of all three proposed techniques in terms of mean, standard deviation and mim/max parameters and their measurement of speed.

Process	**Mean (ms)**	**St. Dev. (ms)**	**Min (ms)**	**Max (ms)**
Segmentation	875	549	179	2630
Feature Extraction	138	17.7	96.5	181
Classification	39.2	7.50	29.2	56.2

Segmentation is the longest stage of the algorithm by far. This is unsurprising due to the iterative nature of the MATLAB GMFIT function i.e an image might go through as many as 100 iterations before a suitable model is found. Additionally, segmentation has the highest time variability of any stage, with times on the given sample ranging from 0.1 to 2.5 seconds. This is due again to the iterative procedure of the Gaussian fitting process i.e an image with “weak” Gaussian clusters will take more iterations to find an accurate model, while an image with strong clusters will take much less time. Adding up the average time for each step, the average time from image capture to the completion of image classification is approximately 1.002 seconds.

4.1 Performance analysis and its Accuracy Analysis

For the first test, a sample of 103 images of my hand was used (the ‘eval/subject1’ dataset). Using the proposed techniques, the gesture recognition algorithm was run on each image and the accuracy of the overall classification was recorded. The results are shown below by gesture in table 2. After the feature extraction from input images using Bag of features through K means clustering, system has to recognize the input-image. The most important feature of SVM-classifier is to allow the non-linear classification. For the analysis of performance proposed system considered 40% of data for training set, by applying of SVM classifier can predict the gesture image and achieves the good results. While calculation of accuracy rate between these two scenarios we get the reached our goal. By comparing with existing approach, the proposed approach reached the 98.2% of average accuracy.

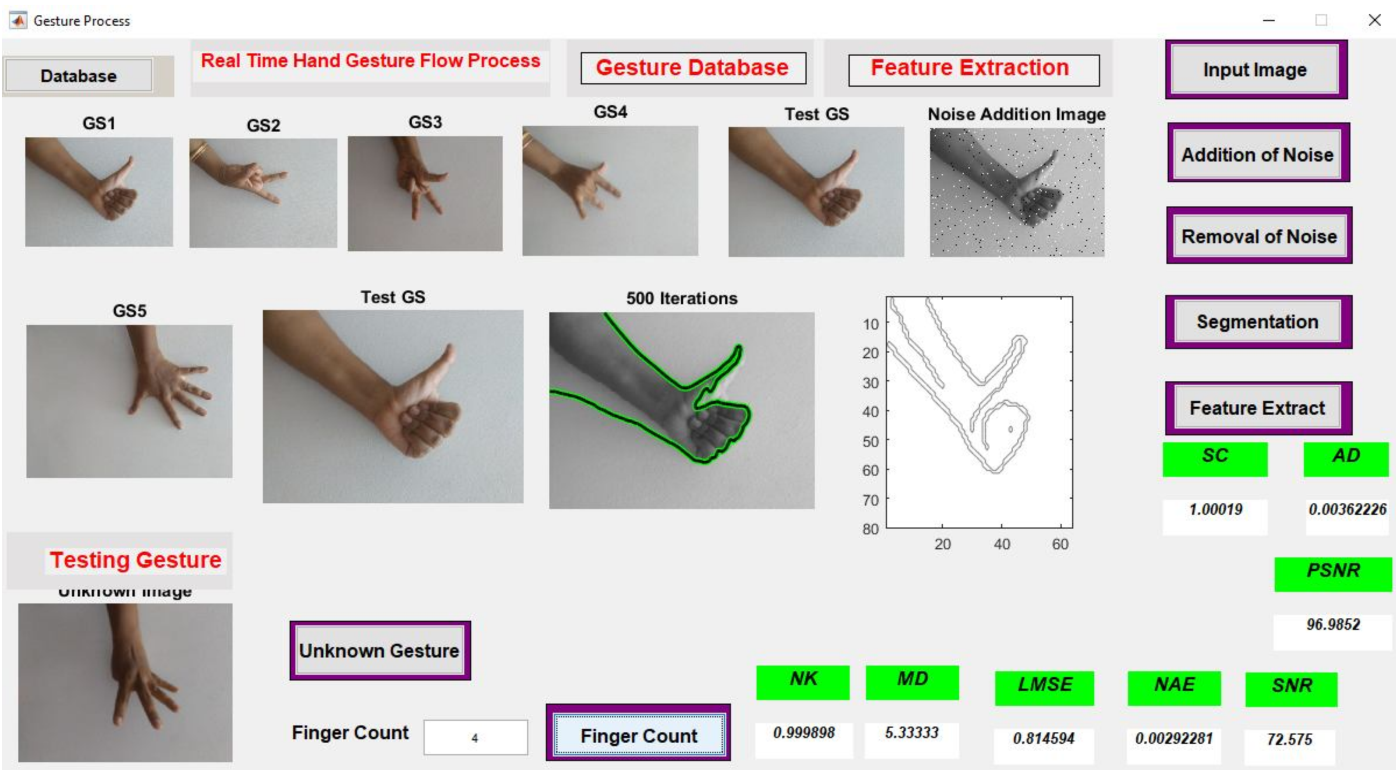

Figure 2(a): Simulated results of proposed hand gesture recognitation through GUI

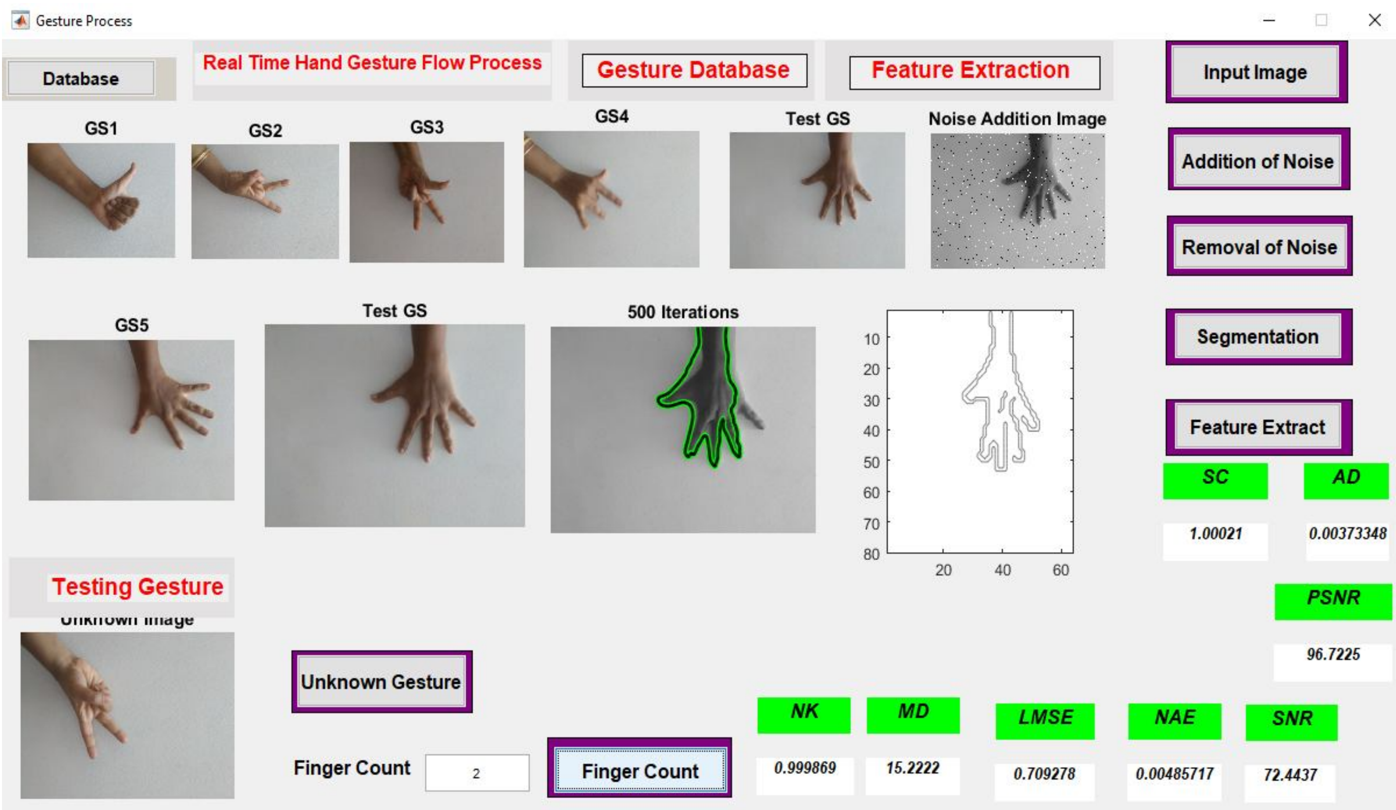

Figure 2(b): Simulated results of proposed hand gesture recognitation through GUI

Table 2: Performance analysis in terms of SNR, PNSR, MSE, AD, SC,NK, MD and NAE for the data base A1-A5 and B1-B5

Parameter /Database	A1	A2	A3	A4	A5	B1	B2	B3	B4	B5
SNR	71.86577	72.17277	72.20389	72.20389	71.81226	73.69608	72.44366	72.73697	72.3755	72.66224
PSNR	95.56674	96.18075	96.24299	96.24299	95.45972	99.22737	96.72251	97.30915	96.5862	97.15968
MSE	0.004265	0.003974	0.003945	0.003945	0.004318	0.002798	0.003733	0.00349	0.003793	0.00355
AD	0.004265	0.003974	0.003945	0.003945	0.004318	0.002798	0.003733	0.00349	0.003793	0.00355
SC	1.000282	1.000328	1.00035	1.00035	1.000289	1.000288	1.000215	1.000261	1.000304	1.000254
NK	0.999829	0.999789	0.999791	0.999791	0.999815	0.999817	0.999869	0.999831	0.999797	0.999838
MD	12.44444	13.44444	14.44444	14.44444	15.91352	13.55556	15.22222	11.66667	15.11111	13.55556
NAE	0.005608	0.006709	0.005886	0.005886	0.006251	0.00629	0.004857	0.006343	0.007153	0.005905

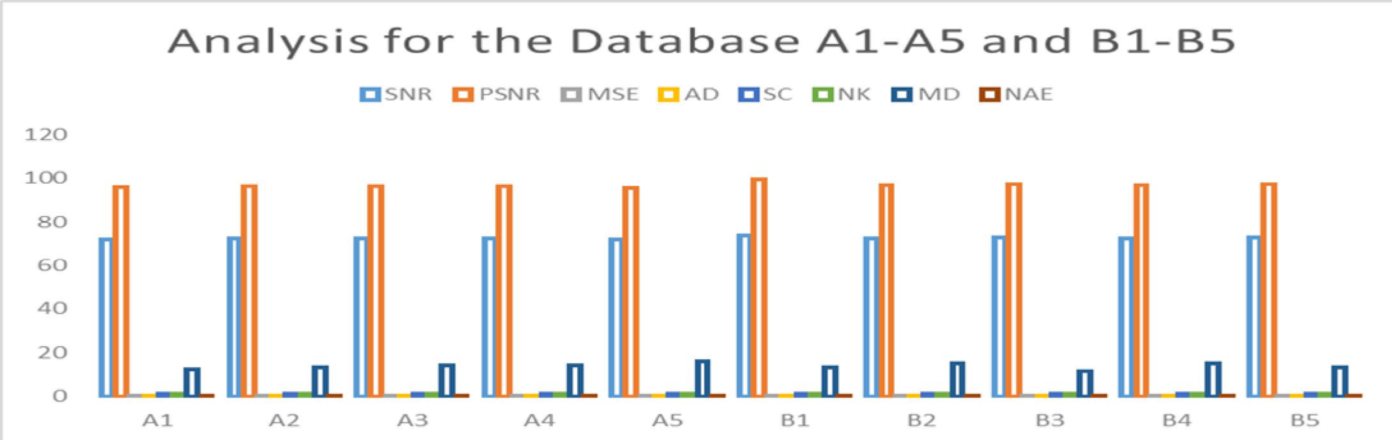

Fig.3.Performance graph for the data base A1-A5 and B1-B5

Table 3: Performance analysis in terms of SNR, PNSR, MSE, AD, SC,NK, MD and NAE for the data base C1-C5 and D1-D5

Parameter /Database	C1	C2	C3	C4	C5	D1	D2	D3	D4	D5
SNR	72.93962	72.74089	71.93244	72.87836	72.03671	72.91901	72.575	73.06058	72.47471	72.8021
PSNR	97.71443	97.31698	95.70009	97.59192	95.90861	97.67323	96.98521	97.95637	96.78462	97.43939
MSE	0.003331	0.003487	0.0042	0.003378	0.0041	0.003346	0.003622	0.003239	0.003707	0.003438
AD	0.003331	0.003487	0.0042	0.003378	0.0041	0.003346	0.003622	0.003239	0.003707	0.003438
SC	1.000199	1.000192	1.00021	1.000285	1.000193	1.000135	1.000188	1.000281	1.000332	1.000189
NK	0.999882	0.999888	0.999876	0.999817	0.999888	0.999922	0.999898	0.99983	0.999796	0.999891
MD	8	8	10.11111	13.22222	10.22222	8.777778	5.333333	9	10.55556	9.563178
NAE	0.00436	0.004093	0.004378	0.006363	0.003962	0.003288	0.002923	0.00549	0.006104	0.003823

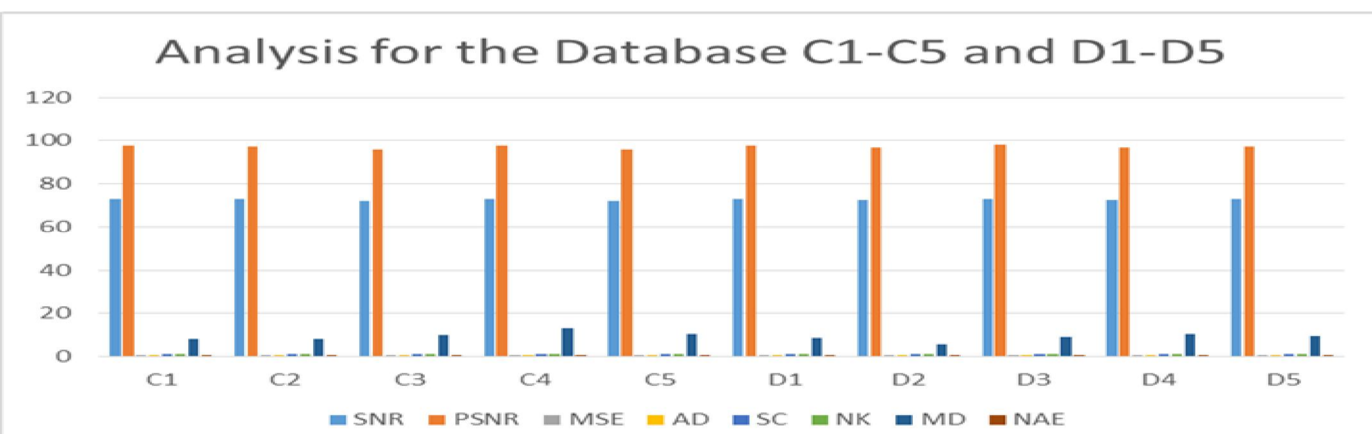

Figure 4: Performance graph for the data base C1-C5 and D1-D5

Table 4: Performance analysis in terms of SNR, PNSR, MSE, AD, SC,NK, MD and NAE for the data base E1-E5 and F1-F5

Parameter /Database	E1	E2	E3	E4	E5	F1	F2	F3	F4	F5
SNR	71.83602	72.0896	71.63602	72.47341	72.47341	54.64038	54.55108	55.06005	55.08113	55.25442
PSNR	95.50723	96.01441	95.10725	96.78201	96.78201	61.11596	60.93736	61.95531	61.99745	62.34405
MSE	0.004294	0.004051	0.004497	0.003708	0.003708	0.225134	0.229812	0.204397	0.203408	0.195451
AD	0.004294	0.004051	0.004497	0.003708	0.003708	0.225134	0.229812	0.204397	0.203408	0.195451
SC	1.000201	1.000165	1.000295	1.000166	1.000166	1.003822	1.003886	1.003828	1.00382	1.003859
NK	0.999886	0.999908	0.999816	0.999906	0.999906	0.997941	0.997902	0.997935	0.99794	0.997879
MD	10.33333	8.222222	13.36225	11.0175	11.0175	39.22222	38.67796	33.33333	37	42.33333
NAE	0.003598	0.002982	0.005741	0.003379	0.003379	0.008905	0.009129	0.009183	0.00922	0.011525

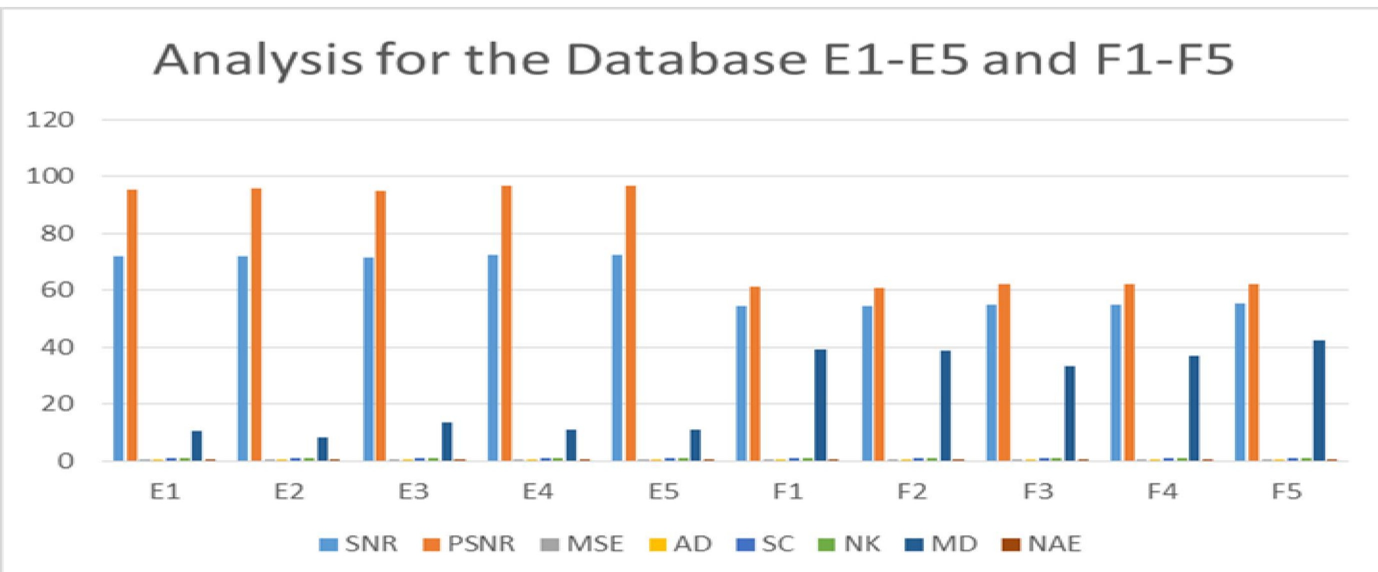

Figure 5: Performance graph for the data base E1-E5 and F1-F5

The performance analysis of proposed hand gesture recognition for the data base A, B, C, D, E and F are applied for proposed methodologies and each has 19, 25,34,28,30 and 23. The table 3, 4 and 5 are shows the parametric analysis in terms of the following.

SNR (Signal to Noise Ration)
MSE (Mean Squared Error)
PSNR (Peak Signal / Noise Ratio)
AD (Average Difference)
SC (Structural Content)
NK (Normalized Cross-Correlation)
MD (Maximum Difference)
LMSE (Laplacian Mean Squared Error)
NAE (Normalized Absolute Error)

The Fig.3 depicts the parametric graph for above said parameters for A1-A5 and B1-B5, Fig.4 for C1-C5 and D1-D5 and Fig.5 is for E1-E5 and F1-F5. The all above paramters are calculated between original hand images and segmented images and from the obtained results, it is found that the accuracy of hand recognition is of 95%.
Table.5 Percentage of gestures recognition in terms of accuracy for four possible hand gestures

Hand Gesture	Total number of Images in data set	Images Correct	Accuracy
A	19	18	98%
B	25	23	94%
C	34	29	89%
D	28	26	94%
E	30	29	96%
F	23	20	95%
Total	103	86	**95%**

Overall the algorithm achieves an accuracy rate of 88% on average. Clearly, the gesture C classification is the weakest of the four, but gesture B recognition is quite good. One way to improve the classifiers overall would be to train them on additional training data. For the training of these classifiers, I used only a small 60 image dataset of my hand. With additional training data, especially from a variety of different people, the classifier would be much more accurate. To improve the gesture C classification in particular (after gathering more data), it would most likely be worthwhile to add another classifier to the pool that is particularly good at correctly identifying gesture C (even if it is not effective at identifying the other gestures). Then, if this new classifier predicts that the gesture is gesture C, its vote would be given a greater weight. In order to test the “big-picture” effectiveness of the classifier, I gathered a second data set from a second test subject and re-measured the accuracy of the classifier. The results are shown below in table 5. As shown, the classifier performs approximately equally well on a person foreign to the system, a desirable characteristic for a hand gesture recognition system.

5. CONCLUSION

In this research work, we propose a framework for hand gesture recognition which uses Bag of features of Kinect, GMM, SVM-KNN and segmentation. Deep information regarding RGB is obtained from the human skeleton through the Kinect camera which is placed in front of the system. With the help of this data and hand movement, hand alignment is carried out which is then transferred to the cloud. Later, employing BOF feature extraction is done for the data stored in the cloud. By comparing the proposed method with the current approach, the average accuracy of the approach is 88%.

REFERENCES

[1]. A.R. Sarkar, G. Sanyal, and S. Majumder, "Hand gesture recognition systems: a survey", International Journal of Computer Applications, 71.15, 2013.

[2]. N. Neverova, Natalia, "A multi-scale approach to gesture detection and recognition", Proceedings of the IEEE International Conference on Computer Vision Workshops, 2013.

[3]. S. Yang, P. Premaratne, and P. Vial, "Hand gesture recognition: An overview", Broadband Network & Multimedia Technology (IC-BNMT), 5th IEEE International Conference, 2013.

[4]. T. Osunkoya and J-C. Chern, "Gesture-based human-computer-interaction using Kinect for Windows mouse control and Powerpoint presentation”, Department of Mathematics and Computer Science, Chicago State University, Chicago, IL 60628, 2013.

[5]. J. Katkar, Jayshree, "Hand Gesture Recognition and Device Control", Hand, 2017.

[6]. C.Selvarathi.et.al, "Human Computer Interaction using Hand Gesture Recognition ", International Journal of Advanced Trends in Computer Science and Engineering,doi.org/10.30534/ijatcse/2020/106922020, Volume 9 No.2, March -April 2020

[7]. M.C. Thomas and A. P. M. S. Pradeepa,“A comprehensive review on vision based hand gesture recognition technology”, International Journal 2.1, 2014.

[8]. H. Zhouand H. Hu,“A survey-human movement tracking and stroke rehabilitation”, University of Essex, Colchester United Kingdom, 2004.

[9]. M. Turkand M.Kölsch,“Perceptual user interfaces”, Emerging Topics in Computer Vision, Prentice Hall, 2004.

[10]. K.K. Vyas, A. Pareek, and S. Tiwari, "Gesture Recognition and Control", International Journal on Recent and Innovation Trends in Computing and Communication ISSN, Retrieved on 16th August, 2017: 2321-8169.

[11]. D.H. Stefanov, Z. Bien, and W-C. Bang,“The smart house for older persons and persons with physical disabilities: structure, technology arrangements, and perspectives”, IEEE transactions on neural systems and rehabilitation engineering, vol.12.2, 2004: pp.228-250.

[12]. D.C. L-Martin, “Universal grammar and American sign language”, Universal Grammar and American Sign Language, 1991: pp.1-48.

[13]. M.L. McIntire, “The acquisition of American Sign Language hand configurations”, Sign Language Studies, 16.1, 1977: pp.247-266.

[14]. M. Alsheakhali, “Hand gesture recognition system”, Computer Engineering Department, The Islamic University of Gaza, Gaza Strip, Palestine, 2011.

[15]. S.D. Badgujar, “Hand Gesture Recognition System”, International Journal of Scientific and Research Publications, 4.2, 2014.

[16]. J. Suarez and R. R. Murphy,“Hand gesture recognition with depth images: A review”, Ro-man, IEEE, 2012.

[17]. K. Patwardhan, K. Srikrishna, and S.D. Roy, "Hand gesture modelling and recognition involving changing shapes and trajectories, using a predictive eigentracker” Pattern Recognition Letters, vol. 28.3, 2007: pp.329-334.

[18]. S. Mitra and T. Acharya, "Gesture recognition: A survey", IEEE Transactions on Systems, Man, and

Cybernetics, Part C (Applications and Reviews), 37.3, 2007: pp.311-324.

[19]. R.Z. Khan and N. A.Ibraheem, "Comparative study of hand gesture recognition system”, Proc. of International Conference of Advanced Computer Science & Information Technology in Computer Science & Information Technology (CS & IT), vol. 2.no. 3, 2012.

[20]. P. Chavan, T. Ghorpade and P. Padiya, "Indian sign language to forecast text using leap motion sensor and RF classifier," Symposium on Colossal Data Analysis and Networking (CDAN), Indore, 2016: pp. 1-5.

[21]. W. Lu, Z. Tong and J. Chu, "Dynamic Hand Gesture Recognition With Leap Motion Controller," *in IEEE Signal Processing Letters*, vol. 23, no. 9, 2016: pp. 1188-1192.

[22]. Kolla Bhanu Prakash.et.al, "Accurate Hand Gesture Recognition using CNN and RNN Approaches", International Journal of Advanced Trends in Computer Science and Engineering, https://doi.org/10.30534/ijatcse/2020/114932020,

[23]. Calvin Ng and Alvin Chua2, "Training of a deep learning algorithm for quadcopter gesture recognition", International Journal of Advanced Trends in Computer Science and Engineering, doi.org/10.30534/ijatcse/2020/32912020, , Volume 9, No.1, January – February 2020.

[24]. Joko Sutopo.et.al, "Gesture Recognition of Dance using Chain Code and Hidden Markov Model", International Journal of Advanced Trends in Computer Science and Engineering,doi.org/10.30534/ijatcse/2019/85862019, Volume 8, N0.6, November-December, 2019.

[25]. Aaron Don M. Africa.et.al, "A Study on Number Gesture Recognition using Neural Network", International Journal of Advanced Trends in Computer Science and Engineering,https://doi.org/10.30534/ijatcse/2019/1484 2019, Volume 8, No.4, July-August, 2019.

[26]. Zavalko N. A.et.al, "Methods for Collecting and Classifying Data using a Soundlet Bayesian Neural Network", International Journal of Advanced Trends in Computer Science and Engineering, doi.org/10.30534/ijatcse/2020/135942020, Volume 9, No.4, July – August 2020