Intelligent Sign Language Recognition Using Enhanced Fourier Descriptor: A Case of Hausa Sign Language

S. T. Hassan[1], J. A. Abolarinwa[2], C. O. Alenoghena[3], S. A Bala[4], M. David[5] and Ali Farizamin[6]
Department of Telecommunication Engineering, School of Engineering and Engineering Technology, Federal University of Technology Minna, Nigeria and Faculty of Engineering, Universiti Malasysia[6.]
salhasst1111@gmail.com[1] j.abolarinwa@futminna.edu.ng[2] carol@futminna.edu.ng[3]
bala.salihu.a@ieee.org[4]
mikeforheaven@futminna.edu.ng[5]

***Abstract*— Hausa sign language (HSL) is the main communication medium among deaf-mute Hausas in northern Nigeria. HSL is so unique that a deaf- mute individual from other part of the country can rarely understand it. HSL includes static and dynamic hand gesture recognitions. In this paper we present an intelligent recognition of static, manual and non-manual HSL using an enhanced Fourier descriptor. A Red Green Blue (RGB) digital camera was used for image acquisition and Fourier descriptor was used for features extraction. The features extracted chosen manually and fed into artificial neural network (ANN) which was used for classification. Thereafter particle swarm optimization algorithm (PSO) was used to optimize the features based on their fitness in order to obtain high recognition accuracy. The optimized features selected gave a higher recognition accuracy of 90.5% compared to the manually selected features that gave 74.8% accuracy. High average recognition accuracy was achieved; hence, intelligent recognition of HSL was successful.**

Key words: Hausa Sign Language; Fourier Descriptor; Particle Swarm Optimization Algorithm; Artificial Neural Network.

I. INTRODUCTION

Gesture is a sign that is either manual or non-manual, and it is meaningful; it is a means of communication that is visual and not vocal. Sign language is a well-coordinated gesture that a hearing impaired or deaf-mute community use to communicate among themselves. Sign language varies from one culture, country to another against the popular knowledge. According to [1], sign language originated from France and many other country's sign languages developed from French sign language. There are different sign languages across the world; in [2] the author introduced American Sign Language, Chinese Sign Language was presented in [3], Arabic Sign Language was introduced in [4], Indian Sign Language was presented in [5], and to mention a few. In Nigeria, Sign language varies across the region, Hausa sign language is the most unique and very difficult for the deaf-mute from other region to comprehend, according to statistics, no fewer than 360 million people across the world are living with hearing impairment [6]. In Nigeria, 8.5 million people are living with hearing impairment and the larger percentage are Hausas, The use of sign language as a means of communication is mostly limited to the deaf-mute community and the deaf-mute from another region or country often find it difficult to understanding a sign of another dialect, hence there is a need for interpreter or translator which is costly and not timing, although so many researchers have been working on different local dialect, therefore it is important to develop an intelligent recognition system for effective communication between the hearing impaired and the hearings locally. There is no existing work on recognition or translation of HSL, this paper present intelligent recognition of HSL using features from Fourier descriptor and also optimizing the features with PSO using ANN for classification which gives us results of accuracy of recognition.

The rest of the paper is organized as follows; section 2 presents related work on sign languages, section 3 is on methodology, section four presents' experimental results and analysis ,and finally, section 5 presents conclusion and recommendation.

II. SIGN LANGUAGE RECOGNITION OVERVIEW

A lot of research work has been done on sign language recognition and so many methodology and different recognition accuracies have been recorded. An instance is the Indian local sign languages [7], owing to the difference in grammatical rule of each language, using the same methodology for different sign language gave different recognition accuracy. Currently, there are several techniques that are applicable for hand gesture recognition with the invention of sensors. In [3], the authors developed an ARM9 sensor data glove [8] that will be connected to computer; it was applied on Chinese sign language recognition. They developed the hardware and the underlying software for data processing and feature extraction for recognition, embedded in the gloves. The approach was able to obtain a good accuracy but it is not natural to user or user's friendly, not cost effective and not error free.

Another recent approach is the Leap motion controller sensor. A small device that is connected to a computer through universal serial bus (USB), It has high accuracy and low cost unlike Microsoft Kinect sensor. Leap motion controller sensor [9] is used to capture data in series of snapshot called frames; each frame contains measured position, velocity and other information from the finger. [4] Worked on Arabic sign language recognition system that translates Arabic sign to test using leap motion controller sensor, the features extracted were fed into classifier for recognition. It was noted that non-manual sign language was not adequately covered; hence, it was recommended that other sensor should be used with leap motion controller so as to obtain higher and better recognition. Microsoft Kinect sensor is a device that is used to capture features from images of signs made by a signer and it is very effective for non-manual signs, [10] used Microsoft Kinect and leap motion controller to capture features from signers, the results gotten from the two features when combined and fed into classifier were better than when the features were used independently. The work was able to complement the limitation of leap motion controller by combining Microsoft Kinect sensor with leap motion controller. Apart from sensor based recognition system, many works has also been done using vision based approach [11]. [12] Worked on Bengali sign language recognition system based on fingertips finder algorithm, the paper proposed a new algorithm for automatic Bengali sign language recognition system. Vision based approached was used; images were captured and eleven features were extracted after preprocessing and segmentation [7], and fed into a multilayer feed forward neural network with a back propagation training algorithm, the finger finder algorithm was able to record 91% accuracy in finger tips recognition.

Another important and fundamental stage in recognition process is feature descriptor. Feature descriptor can be either discontinuity based or similarity based. Discontinuity or edge based is mostly used for shape description since it is a contour based, that is, edge of the shape. Fourier descriptor is a contour based descriptor; it is invariant to size or scale, translation and rotation which are the primary conditions used to know a good descriptor. There are many works that have used Fourier descriptor to extract features, [13] and [7] used Fourier descriptor and other descriptors to achieve better recognition rate, and [14] also used Fourier descriptor with 7hu moment for features extraction. The next section is on methodology used for intelligent Hausa sign language recognition using enhanced Fourier descriptor.

III. METHODOLOGY

The flow chart below shows the stages of implementation of the work.

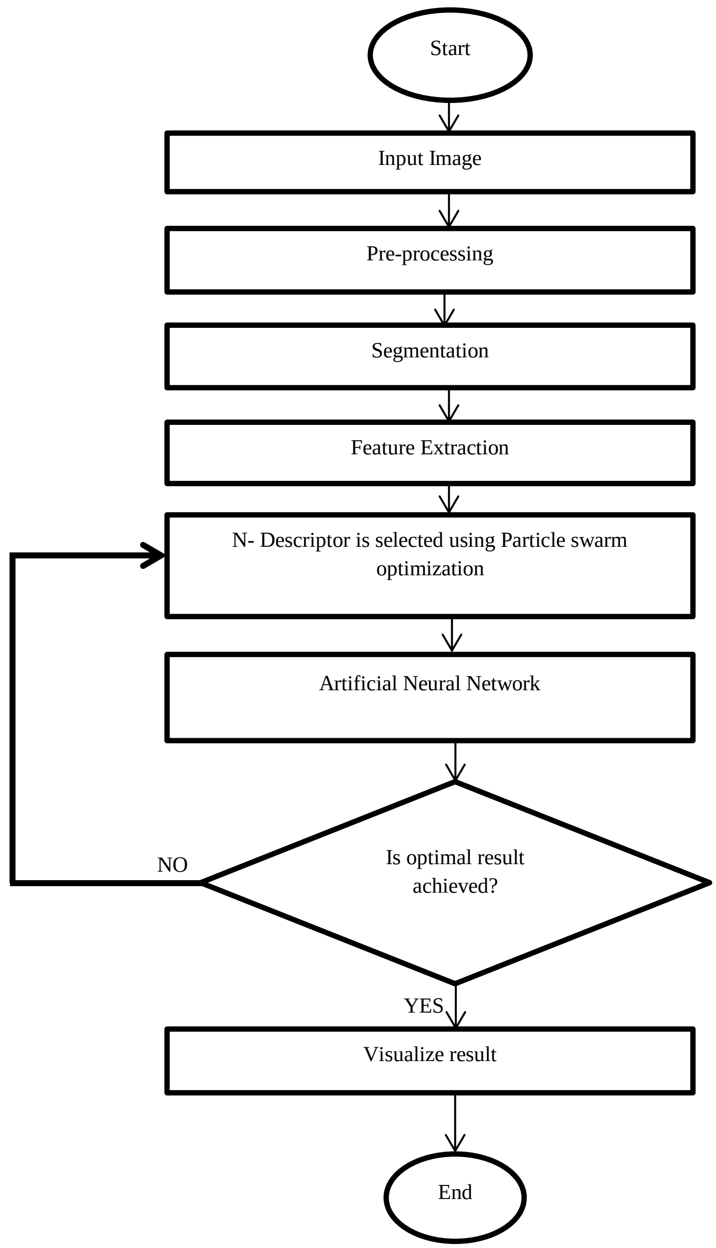


Fig. 1. Flow chart of the methodology

A. Input Images, Preprocessing and Segmentation

There is no source of Images or video standard dataset of HSL. Images were captured using RGB digital camera at varied lighting conditions, the 21 static, manual and non-manual signs were labeled class 1-21 and each of the class has 10 samples. The captured images were resized to a uniform dimension 500X500. The images were converted from RGB images to gray-scale images with varied intensity of 0 to 255, and then, the images were converted to binary images. Thresholding (adaptive) method was used in the said conversion and 7X7 median filtering was used to remove unnecessary noise. After these preprocessing and thresholding, we had segmented images. The Prewitt edge detection was used to detect the edges so as to obtain the boundaries coordinates.

B. Feature Extraction

For the features extraction, a contour or discontinuity based shape descriptor is needed, hence Fourier descriptor was used because of its advantage over other descriptors. The Discrete Fourier Transform of the contour pixels is what made up of the feature vector which was gotten with the mathematical illustration in (1) and (2).

$$S(K) = [x_k + jy_k] \qquad (1)$$

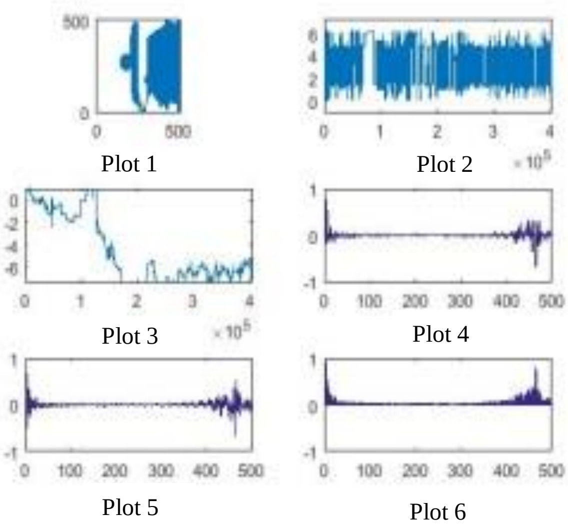
$$k = 0,1,2,\dots,k-1$$

$$a(u) = \sum_{k=0}^{k-1} S(k) e^{-j2\pi uk}/_{K} \qquad (2)$$

$$u = 0,1,2,\dots,k-1.$$

Where S(K) is the complex coordinate of the boundary pixels, K is the total number of pixels in the image, a(u) complex coefficient of the Fourier descriptor of the boundary coordinates [12], the information compactness of DFT and its rotation, translation and scale invariance properties added to the accuracy rate of the transformation. The Fourier transformation produce a complex output, the magnitude and the phase value, both are needed to preserve the information but most information is available at the magnitude of the spectrum, however ,since no need for inverse transformation, only the magnitude value of the data was be considered. The table below is a sample of the features extracted using DFT and the plots of all the important features were shown in table 1.

TABLE I. A TABLE FOR A SAMPLE SIGN AND THE FOURIER DESCRIPTORS PLOTS

A Sample Sign	Feature Descriptors Plots Obtained From MATLAB	Names of the plots
		Plot 1: Curvature plot. Plot 2: Phase plot. Plot 3: Cummulative Phase plot. Plot 4: Normalized Magnitude plot. Plot 5: Fourier Coefficients plot. Plot 6: Angular Coefficients plot.

C. Application of Particle Swarm Optimization

The Particle Swarm Optimization algorithm was used for the selection of the descriptors and it will check for best fitted global descriptors among all the swarm of the candidate solutions, the algorithm will ensure that the positions of the candidate solutions and the search for the best fitness are in order and keep memory of the previous event and that will enable the classifier to have highest and best recognition accuracy. The mathematical equation for the algorithm is as shown below:

$x_i =$ Candidates' solution (the descriptors gotten using Fourier descriptor).

$$x_i = \{x_1, x_2, x_3, x_4, \dots, x_D\}$$

$$x_i(t+1) = x_i(t) + v_i(t+1) \quad (3)$$

D. Artificial Neural Network

An artificial neural network is a computational model inspired by the neural structure of human brain. A feed forward neural network in combination with a supervised learning scenario is used in this work. It is a back propagation algorithm. The input patterns are given to the network through the neurons in the input layer and the output of the network is obtained through the neurons in the output layer. Training Phase; the neural network is trained to classify 21 static, manual and non-manual HSL. The training dataset contains 70% of the total data set. Testing Phase; in this phase, 15% of the dataset containing all the classes and the samples were used. Lastly, the Validation phase uses the remaining percentage of the dataset. Below are mathematical equations and explanations that illustrate the training and testing processes;

1. *Activation Function*

It translates the input signals to output signals when it receives output from the summation processor. The activation function uses a threshold to produce an output. There are four types of activation functions. In this research, sigmoid function was employed as stated in (4).

$$f(x) = \frac{1}{1+e^{-bx}} \quad (4)$$

Where: $f(x) =$ Output signal, x = input signal or descriptor,

1. *Training Procedure*

a) *Step 1:* for each neuron in the hidden layer, compute the weighted inputs using this equation 5,

$$U_k = \sum_{j=1}^{D} w_{kj} x_j + b_k \quad (5)$$

Where; $x_j =$ input signal or descriptor, $w_{kj} =$ weight from the jth to kth neuron and b represent the bias.

$$v_i(t+1) = v_i(t) + c_1\big(p - x_i(t)\big)r_i + c_2\big(g - x_i(t)\big)r_2 \quad (6)$$

Where $v_i(t+1)$ is the vector collecting the velocity component of the ith particle along the D-dimensions, $v_i(t)$ is the inertial components, $c_2\big(p - x_i(t)\big)r_i$ is the cognitive component, $c_1\big(g - x_i(t)\big)r_2$ is the social component, p is the personal best of the descriptor and g is the global or overall best of the descriptors.

b) *Step 2*: Applying the activation function to the results obtained in step 1 to get the outputs of the neurons in the hidden layer.

$$Z_k = \mu(U_k) = \frac{1}{1 + e^{-\sum_{j=1}^{D} w_{ki}^1 + b_{ki}}} \quad (7)$$

μ Represents activation function and Z_k represents output of the neurons.

c) *Step 3*: Then each neuron in the output layer is the same as the number of the learning problem, the computation of their weighted input was determined using mathematical model given by

$$Y_p = \sum_{j=1}^{D} w_{pi}^2 z_j + b_{p2} \quad (8)$$

Where p is the number of the output neurons.

d) *Step 4*: Applying the activation function to the results gotten from step 3 to obtain the output of the neurons in the hidden layer, it becomes:

$$Y_p = \mu(U_k) = \frac{1}{1 + e^{-\sum_{j=1}^{D} w_{ki}^1 + b_{p2}}} \quad (9)$$

e) *Step 5:* Final classification of the ANN is the neuron that contains the maximum output Y_p which indicates classification.

IV. RESULTS AND ANALYSIS

This section presents the implementation results of the intelligent Hausa Sign Language recognition system. The system was implemented on MATLAB R2015b on a personal computer. Experiments were conducted for 21 classes with 10 samples each for the static, manual and non-manual signs. 70% of the dataset was used for training, 15% for testing and 15% for validation. The simulation ran without optimization, and an average recognition accuracy rate of 74.8% was recorded. Fig. 2

shows the Cross-Entropy plot showing best validation performance without optimization at 0.073834 at epoch 28.
.

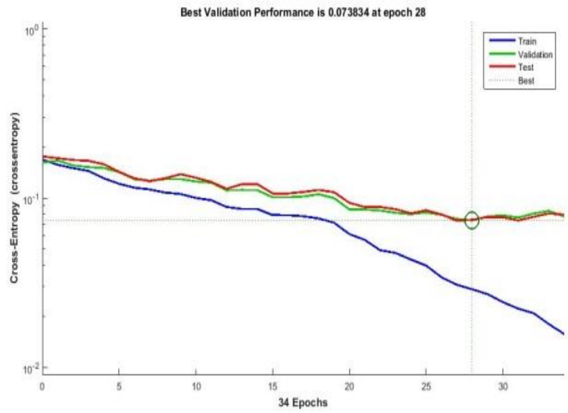


Fig. 2. Cross-Entropy Plot Without optimization

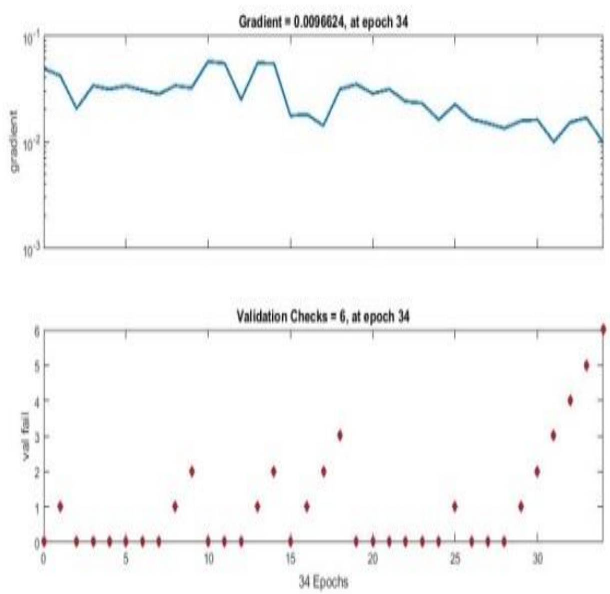


Fig. 3. Validation Fail check and gradient plots without optimization

When the simulation ran with PSO algorithm to enhance the descriptors, an average accuracy rate of 90.5% was achieved. Fig. 4 shows the Cross-Entropy plot showing the best validation performance of 0.06141 at epoch 65. While Fig. 5 shows the validation failure check. The gradient plot is close to horizontal axis; hence, its value of 0.0024068 at epoch 71 is lower compared to that of the un-optimized gradient value of 0.096624 at epoch of 31 shown in Fig. 3. This shows how accurate the result obtained when optimized is.

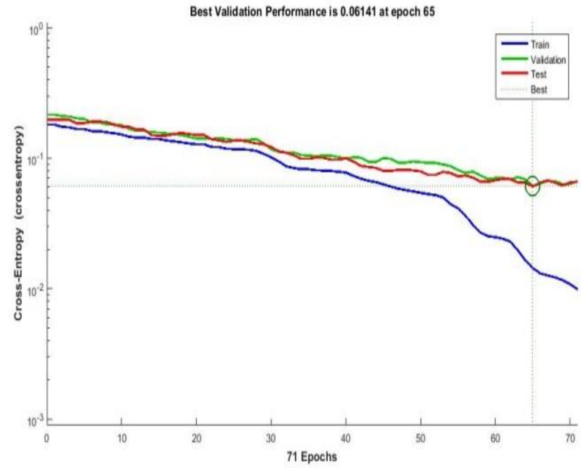


Fig. 4. Cross-Entropy Plot When Optimized

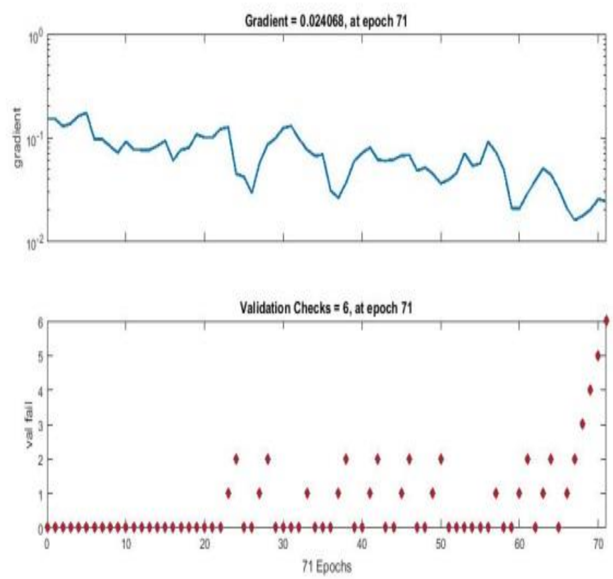


Fig.5. Validation check and gradient plots when optimized

V. CONCLUSION

Communication between the hearing impaired and the hearings is difficult without translator. There is no existing work on HSL recognition system. This work was able to develop intelligent recognition system for HSL using features obtained from Fourier descriptor. The features were optimized using ANN in order to check the fitness of the features. The average recognition accuracy obtained when optimized was impressive, thus, we have succeeded in developing intelligent recognition of HSL which in turn annihilate the barrier of communication between the deaf-mute and the hearings. It is recommended that consideration be given to dynamic signs in future work.

REFERENCES

[1] M.F Tolba, "Recents development in sign language recognition system," IIEEE, pp. xxxvi-xlii, 2013.

[2] C. Dong, "American sign language alphabet recognition using microsloft kinect," IEEE, pp. 44-50, 2015.

[3] Li Lei and Que Dashun, "Design of data-glove and Chinese sign language recognition system based on ARM9," IEEE, pp. 1130-1135, 2015.

[4] A.S.Elons, "Arabic sign language recognition using leap motion sensor," IEEE, pp. 330-335, 2014.

[5] A. S. Ghotkar, "Vision based multi-feature hand gesture recognition for Indian sign language manual," International journal on smart sensing and intelligent systems vol. 9, NO. 1, pp. 124-147, 2016.

[6] Abdulazeez Ahmed et al, " Hearing impairment in asemi-urban community in north-west Nigeria," European journal of preventive medicine, pp.113-119, May,2006 .

[7] G. K. Archana and S. Ghotkar, "Hand segmentation techniques for hand gesture recognition," International journal of human computer interaction (IJHCI), pp. 15 - 25, 2012.

[8] G. Saggio, "A novel array of flex sensor for goniometric glove," Elsevier, pp. 119-126, November 2013.

[9] Jayash Kumar Sharma, Rajeev Gupta and Vinay Kumar, "Numerical gesture recognition using leep motion sensor," Internaltionl conference on computation inteligent and communication Network ,IEEE , pp. 411-413, 2015.

[10] Giulio Marin, Fabio Dominio and Pietro Zanuttigh, "Hand gesture recognition with jointly calibrated leap motion and depth sensor," Springer science and business media new york, pp. 14991-15015, 2016.

[11] Siddharth S. Rautaray and Ampanm Argrawal, "Vision based hand gesture recognition for human computer interaction:a survay," Spring science and business media, pp. 1-54, November 2012.

[12] J. Angur and M. Jarman, "An automated Bengali sign language recognition system based on fingertip finder algorithm," International journal of electronics & informatics, pp. 1-9, 2015.

[13] G. Adithya V., "Artificial neural network based method for Indian sign language recognition ," Proceedings of 2013 IEEE Conference on Information and Communication Technologies (ICT 2013), India, 2013.

[14] Peng Jinye and Yang Quan, "Chinese sign language recognition for a vision-based multi-features classifier," International Symposium on computer Science and computational technology, IEEE, China, pp.194-197,2008.