

NEURAL NETWORK RECOGNITION SYSTEM FOR IGBO VOWELS

Ebenezer Obaloluwa Olaniyi¹, Oyebade Kayode Oyedotun², Khashman Adnan³

^{1,2}Near East University,

Lefkosa, Mersin 10, Turkey,

Member, Centre of Innovation for Artificial Intelligence
obalolu117@yahoo.com, oyebade.oyedotun@yahoo.com

³Founding Director, Centre of Innovation for Artificial Intelligence

British University of Nicosia

Girne, Merin 10, Turkey

adnan.khashman@bun.edu.tr

Abstract—Handwritten recognition is one of the problems facing researchers in the field of computer vision that many researchers are still finding it difficult to provide solution. Several researches have been carried out on English Language and Asian Character recognition with little work done on the African language character recognition. In this paper, we consider Igbo vowels character recognition; we employed several people to write Igbo vowels character with their handwritten. These characters differ in their shapes, size, and orientation. These characters are then resampled by rotating each one of them, this make the system to be more intelligent. In this work, the process of recognizing the Igbo vowel character was divided into two stages. In the first stage, preprocessing of the characters was carried out by converting the characters into binary using the threshold value obtained from the Otsu's method. After that noise was removed by filtering. In the second stage, the feature extraction was carried out to bring out the important features of the image. These were further resized and reshaped to be fed into the neural network. The multilayer feedforward neural network was created and trained using backpropagation algorithms. After the training has been carried out, the network was simulated. The result obtained from the simulation gave a recognition rate of 90.2%.

Index Terms—Igbo Vowels Character, neural network, recognition, backpropagation.

1 INTRODUCTION

Character recognition has been one of most difficult problem due to the variation in the way different people write, different sizes and also orientation of an individual handwritten varies.

In Character recognition, printed documents are transformed into ASCII files for the purpose of editing, compact storage, fast retrieval through the computer[1]. Character recognition can be divided into two: (1) Online character recognition in which text is automatically converted as it is written on digitizer such as pc tablet, where a sensor picks up the pen velocity as characters are scripted. The signals obtained are transformed into a letter code which is usable to a computer and text processing application. The second type is offline character in which handwritten characters are scanned in form of paper document, process and converted to binary or gray scale to make available to a recognition algorithm.

Many research works have been done on Chinese[2] with recognition rate that is more than 98%, English character with recognition rate of 92.5%[3], English Vowel Character[4], Arabic character with recognition rate of 97%[5] and Farsi Language with recognition rate of 85%[6] but little work has been done on African character. Character recognition has many applications. In today's technology, Google translate apply character recognition. It recognized characters based on the previous learning by the system. In this work, Igbo vowel characters have been considered because of its importance in

the southern eastern of Nigeria. Igbo language is one of the three major languages in Nigeria; it is widely spoken by approximately 18 million people in the southeastern part of Nigeria and also recognized by minority in Equitorial Guinea[7]. It is widely speak to carrying out daily today activities, it is speaks in customary court, churches, markets, festival activities. Therefore, it is important to develop an automatic vowel recognition system for Igbo language. In this work, we exploited neural network for Igbo vowel handwritten character recognition. Neural network has widely been used in character recognition.

In this paper, we present multilayer feedforward network, trained with backpropagation algorithm and develop an intelligent recognition system using neural network.

The rest of the paper were arranged as following: the section two(2), explain a brief explanation on the Igbo vowels, section three(3) explain Artificial neural network and its operational principles, section four(4) explain the result and analysis, section five(5) is the conclusion.

2 IGBO LANGUAGE

The Igbo can also be referred to as Ibo. Igbo people are one of the largest single ethnic groups in Africa. Most of the Igbo speakers dominate Southeastern part of Nigeria. They made of 17% population of Nigeria[8]. They can also be found in signif-

icant number in Cameroon and Equatorial Guinea. The language spoken by this people is called Igbo. The major Igbo people in Nigeria have their settlement in the following state; Anambah State, Imo State, Abia State, Ebonyi State, and Enugu State. Also, the Igbo people takes about 25% population of some other state in Nigeria like Delta state, and River State. The traces of Igbo culture and language can also be found in the south southern and south west of Nigeria, in state like Lagos, Rivers, Akwa-Ibom, Bayelsa and Ogun state. Igbo language has 36 alphabet; nineteen(19) consonant, eight(8) vowels and nine(9) are diagraph.

The Igbo vowels are written below with draw pad software:

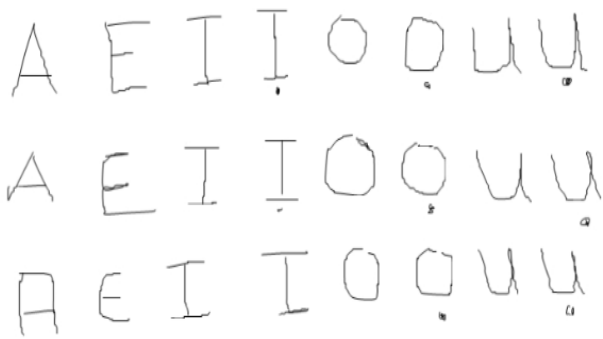


Fig1. Handwritten of Igbo vowels with draw pad.

2.1 PREPROCESSING

The handwritten vowel characters were obtained by asking friends to write the vowels on the drawing pad software. These were later saved as a picture using Jpeg format. The total number of saved handwritten vowel is 80. This folder was then read into matlab for processing. Some of common processing operations performed before recognition are: binarization, morphological operation, edge detection, dilation, skeletonization, rotation of images and feature extractions.

2.2 BINARIZATION

The images were read into the matlab inform of true colour which need to be converted to binary. The conversion depends on the thresholding value of the gray. Thresholding create binary images from gray level by turning all pixels below some threshold value to zero(background) and all pixels above the threshold value to one(foreground)..There are three types of thresholding (1) Global thresholding (2) Local thresholding(3) Dynamic or adaptive thresholding. Thresholding may be viewed as an operation that involves test against a function T of the form[9]:

$$T = T[x,y p(x,y),f(x,y)].....(1)$$

Where $f(x,y)$ -gray level of the point(x,y)

$P(x,y)$ - some local property of this point

A thresholded image $g(x,y)$ is defined as

$$g(x,y) = 1 \text{ if } f(x,y) > T \text{ or } g(x,y) = 0 \text{ if } f(x,y) \leq T \dots(2)$$

The pixel label 1 correspond to object(foreground) and the pixel label 0 correspond to the background.

Whenever T depend only on $f(x,y)$ i.e. only the gray level values, the threshold is known as Global threshold. But if, T depends on both $f(x,y)$ and $p(x,y)$, the threshold is called local. If

T depend on the spatial coordinate of x and y then the threshold is called dynamic or adaptive threshold.

2.3 MORPHOLOGICAL OPERATION

Morphological operations depend on two argument which are binary images and structuring elements. The binary images are the images with two possible values for each pixel. The two colour used in binary images are black and white which is denoted as (1 and 0). The two colour are the foreground colour and the background colour. The structuring element is a shape used to interact with a given image with purpose of drawing conclusion on how this shape fit or misses the shape in the image. It is typically used in morphological operation such as dilation, erosion, opening and closing or as well as hit or miss transformation. The morphological operation compare the structuring element to neighbourhood of each pixel to determine the output of its operation.

2.3.1 Edge Detection

Edge detection is a fundamental tool of image processing. It is usually employ for detection and extraction of feature in an image. Edge detection aims identifying the point at which there is a sharp changes in the brightness of a digital image and find discontinuities using the amount of d. The main purpose of using edge detection is that it helps in terms of significantly reducing the amount of data in an image and preserve the structural properties for further image processing[10].

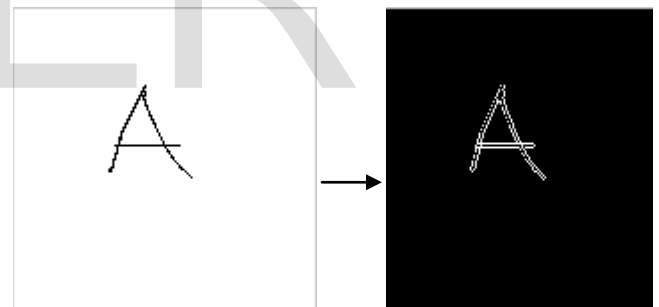


Fig 2 shows conversion of global threshold binary image to edge detection.

2.3.2 Dilation

Dilation is a mathematical morphology operator that is basically used in binary image though there is a version that can also be applied to grayscale images. Dilation works by gradually enlarging the boundaries of the region of the foreground pixels. The foreground areas grow in size while the holes within those region becomes smaller.

Dilation being one of the morphological operator, it has two arguments which are the binary image to be dilated and the structuring element. The structuring element is usually small and it determines the precise effect of the dilation on the image[11]. The structuring element that was applied in this research work was 3*3 square with the origin at the centre. To determine dilation of a binary input image by the structuring element. The background pixels in the input image are consid-

ered one after the other. For each of the background pixels we place the structuring element on top of the input image so that the origin of the structuring element and the input pixel position coincide[11]. The pixel is set to be foreground, if at least one of the pixel in the structuring element correspond with a foreground pixel in the image underneath. But if all corresponding pixel in the image are background, in this case the input pixel is left at the background value. The diagram below shows the process of dilation.

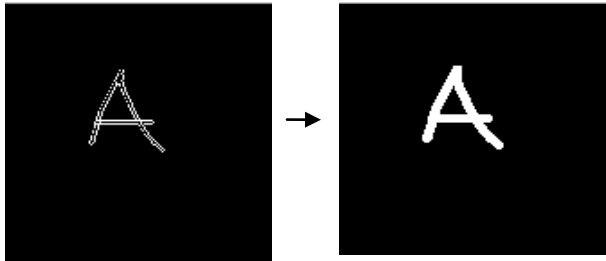


Fig 3 Conversion from edge detection to dilation

2.3.3 SKELETONIZATION

Skeletonization is a morphological operation that provides skeleton of all points belong to the medial axis, but due to pixelization and finite size of the structuring element, it is not necessarily connected nor maximally thin[12]. It reduces the unnecessary pixel from an image and makes the image to left with one pixel thick, it preserve the topology of an object. It makes processing faster and also enhances recognition of images by recognition algorithms. The diagram below is a typical illustration of skeletonization as experiment in this research work.

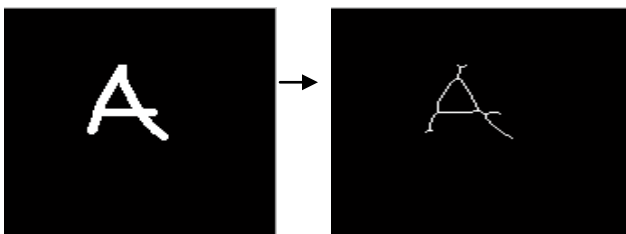


Fig 4 Shows conversion from dilation to skeletonization

2.4 Rotation of the images

Despite the variation in the writing style, shape of an individual handwritten, each handwritten vowels are rotated into two direction, 45 degree and 330 degree was considered in this work. This was done to build into the network some rotational invariance. This makes the network to be more intelligent in recognizing characters with moderate displacements in orientation. The resampling of the image is shown in the figure below



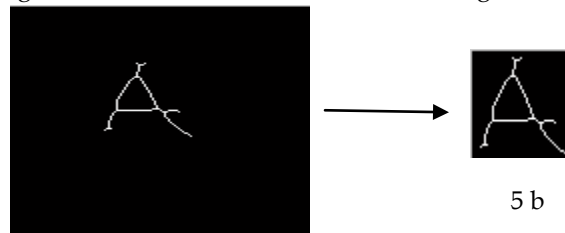
4a

4b

Fig 4a shows the 45 degree rotation of the images, Fig 4b show 330 degree rotation of the images

2.5 FEATURE EXTRACTION

Feature is extracted from the character by cropping the important features. The sub-images have to be cropped sharp to the border of the character in other to standardize the sub-images[13]. The image standardization was done by finding the maximum row and maximum column with 1s and with peak point, increase and decrease the counter until meeting the white space or line with all 0s[13]. The technique of carrying out feature extraction is shown in the figure below



5 a

5 b

Fig. 5a Skeletonization of image, Fig. 5b Cropping of image.

After the image has been pre-processed, The next step is to resize the image to meet up with the network input requirement. There is a need to be careful in resizing an image, as the blurriness of an image should be under consideration so as to make all the extracted feature be clear enough for the network to comprehend. In this work, the image was resize to 40 by 40 matrices which comprises background and foreground. The background is made up of 0s(black) and the foreground is made up of 1s(white) pixels. The image is then reshaped to be fed into the neural network. It is reshaped to 1600 * 1 which was concatenated to 1600 * 240 with binary matrices(1 and 0).

3 ARTIFICIAL NEURAL NETWORK

Artificial neural networks are computational networks which attempt to simulate, in gross manner, the networks of nerve cell(neurons) of biological(human or animal) central nervous system[14]. It allows very simple computational operator such as addition, multiplication to solve non- linear problem. It is a massively connected machine or system that model human brain in two aspects; (1) Function and (2) Structure. In function, networks are modeled after human ability to optimize, think, make decision, forecast, and associate. In structure, it is modeled on biological neurons based on the arrangement of brain neuron such as dendrites which receives input data from the sensor, cell body that activates the neuron, axon which is the transmission channels for electrical signals and synapse which is transmitter of chemical signals. The primary unit of artificial neural network is an artificial neuron. It is a simple mathematical function that has three set of rule which are summation, multiplication and activation(transfer function). At the input of the neuron, the inputs are weighted, which implies that the inputs value are multiplied with indi-

vidual synaptic weight. There is a sum function that sums all weighted inputs and bias. At the exit of neuron the sum of previously weighted inputs and bias is passing through activation function that is refer to as transfer function. In this work, sigmoid function is use as transfer function because of its soft switching and asymptotic properties.

A Back-propagation neural network is a multilayer feed-forward neural network, it considered as one of the simplest and most general method used for supervised training of multilayer neural network[15]. Backpropagation neural network is a supervised learning network that learns by example. If a sample is given to the network as an input there must be corresponding target of the sample. It learns by adjusting the synaptic weight internally. It can generalize input that is not included in the training sample.

3.1 OPERATION OF FEED FORWARD ARTIFICIAL NEURAL NETWORK

Feedforward neural network is arrange structurally as input layer, hidden layer and output layer. The input layer serve as passage for the parameters into the network. It is called non processing layer. The hidden layer consist of numbers of neuron or hidden units that is placed in parallel. It is processing layer. It enhance the network to learn complex task by extracting progressively meaningful features from the input patterns (vectors) [16]. Each neuron performs a weighted summation of the inputs and these are later passed into non linear activation function (σ). The non linear activation function in the neuron is usually chosen to be smooth step function because of its smooth and soft switching.

Mathematically,

$$\text{Sigmoid}(x) = \frac{1}{1+e^{-x}} \dots \dots \dots 3$$

When the input pattern is passed into the neural network via the input neurons. The synaptic weights passes the input pattern into the hidden layers which is the processing layer, the weighted summation of the input is carried out and then passed it to the activation function (sigmoid function). The sigmoid function has a threshold value that is either 1 or 0. If the values produce from the weighted summation that is passed to the sigmoid function is greater or equal to 1. Then an output of one (1) will be produce but if the weighted summation is less than one the output zero (0) will be produce in the hidden layer. The output value of the hidden layer is then passed to the output layer. The output layer is also a processing neuron in which weighted summation of the hidden values is carried out. After this is done by the neuron in the output layer, the result is produce as the actual output.

3.2 OPERATION OF BACK PROPAGATION NEURAL NETWORK

Backpropagation network is a supervised neural network that learns by example. A set of sample is given to the network with its corresponding classes or target. In this research batch training was employed in which all the input training was introduce to the network once with their corresponding target. During training, errors occur due to the difference between the actual output and target output.

The following are the equations of backpropagation[15]
 Output layer

Assuming O_k is the activation function value produce from the output layer of neuron k and the corresponding target of neuron k is t_k .

$$e_k = t_k - O_k \dots \dots \dots (4)$$

The error signal for neuron k in the output layer

$$\Delta k = (t_k - O_k) O_k (1 - O_k) \dots \dots \dots (5)$$

$O_k (1 - O_k)$ is derivation of the sigmoid function.

The weight $w_{j,k}$ between the output node k and node j can be modify by

$$\Delta w_{j,k} = l_r \delta_k x_k \dots \dots \dots (6)$$

$$w_{j,k} = w_{j,k} + \Delta w_{j,k} \dots \dots \dots (7)$$

Where $\Delta w_{j,k}$ is the change weight between the node j and k
 l_r is the learning rate

The equation below is used to improve the process of updating the weight

$$\Delta w_{j,k} = l_r \delta_k x_k + \Delta w_{j,k}^{n-1} \mu \dots \dots \dots (8)$$

μ is the momentum rate. This was introduced to accelerate the learning process by enhancing the learning process to continue in the same direction with larger steps.

HIDDEN LAYERS

The error signal for node j in the hidden layer

$$\Delta k = (t_k - O_k) O_k \sum (w_{j,k} \delta_k) \dots \dots \dots (9)$$

To adjust the weight between the $w_{i,j}$ input neuron i and neuron is

$$\Delta w_{i,j} = l_r \delta_j x_k + \Delta w_{i,j}^{n-1} \mu \dots \dots \dots (10)$$

$$w_{i,j} = w_{i,j} + \Delta w_{i,j} \dots \dots \dots (11)$$

3.3 DESIGN OF THE ARTIFICIAL NEURAL NETWORK

The dataset that we used in this research work was collected from ten different people who we employ to help in using wordpad to write the Igbo vowel. The samples were later resample by rotating each alphabet into 45-degree and 330-degree. After all the processing of the image was carried out and size of image was resize to matrix [40 40]. This was later reshape to matrix [1600 1] in order to be fed into the neural network using neural network tool box in Matlab (NNTOOL). The total dataset was sum up to 240 samples. This later was divided into training set and testing set. 70:30 ratios was used to divide the dataset which gives 168 dataset for training set and 72 dataset for the testing set. The reshape input matrix of 1600*168 training with corresponding target 8* 168 matrix was fed into the neural network since NNTOOL use batch training. The number of input neuron in the input layer is 1600 neurons and the hidden layer neuron used was six (6), this hidden neuron was experiment from two (2) to Eight (8) and six (6) neurons was discovered to be the best hidden neuron for this research because it prevent under fitting and over fitting of the training. Also, it prevents the network to stuck at the local minima. The output layer is made up of eight neurons which represent each vowels.

The figure below illustrates the design of the vowel recognition neural network.

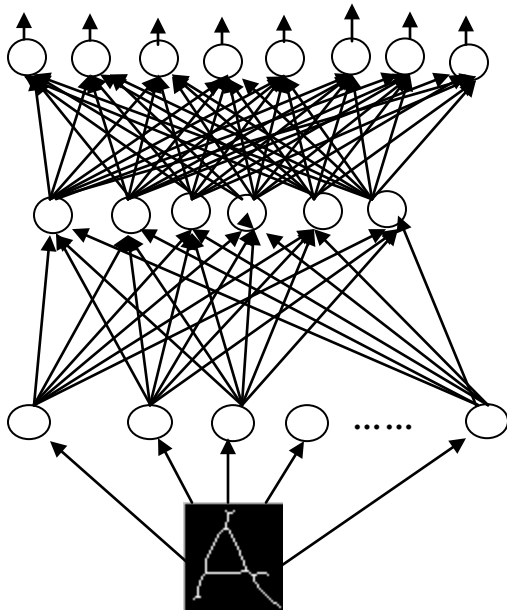


Fig 6: Architecture of feedforward neural network.

4. RESULT

The training was carried out by adjusting the learning rate and the momentum rate. The learning rate determines the speed at which the network learns from the input samples while the momentum rate is used to reduce the possibility of the network from settling at a local minima. Also, it helps the learning process by encouraging the changes in weight to continue in the same direction with larger steps. The learning rate that was finally achieved and made the network to converge is 0.31 with the momentum rate of 0.72. It took the network 27 seconds to be trained and gave a recognition rate of 90.2% on testing. The table and the graphical representation below displays the result obtained for handwritten vowel recognition using artificial neural network.

Table1: PERFORMANCE TABLE OF NEURAL NETWORK

No of training sample	168	No of Input neuron	1600
No of testing sample	72	No of hidden neuron	6
No of correct recognition	65	No of output neuron	8
Learning rate	0.31	Time spent	27sec.
Momentum rate	0.72	Performance	0.2398
Recognition rate	90.2%	Epoch	2000

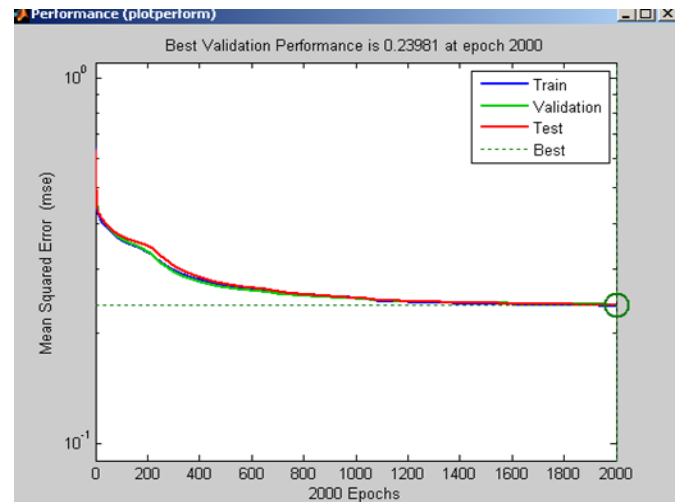


Fig 6: Graph of minimum square error against Epoch

5. CONCLUSION

This work proposed a system for recognizing Igbo vowel character using artificial neural network.

In this research work, recognition rate of 90.2% was obtained after testing the neural network with the 30% of the dataset. This algorithm has shown that despite the variation in the handwritten of an individual in terms of size, shapes and orientation, artificial neural network system can recognize and classify accordingly.

In future works, it will be interesting to carry out research on all the thirty six Igbo character in which the authors are also working on at present.

REFERENCE

- [1] Guarav Kumar, PradeepBhatia, "Neural network based Approach for Recognition", In International Journal of Computer Application, Vol. 62, 2013, pp. 8-13.
- [2] LiangbinZheng, Ruqi Chen, XiaojinCheng, "Research on Offline Handwritten Chinese Character Recognition Based on BP Neural Network" In International Conference of Computer Science and Information Technology, Vol. 51, 2012, pp.323-329
- [3] Tirtharaj Dash, TannisthaNayak, "English Character Recognition Using Artificial Neural Network" In Conference of AIREs, 2002, pp. 7-9.
- [4] Manish Mangal, Manu Pratap Singh, "Handwritten English Vowels Recognition Using Hybrid Evolutionary Feed-Forward Neural Network", Malaysian Journal of Computer Science, Vol. 19(2), 2006, pp 169-187.
- [5] Haithem Al-Ani Ban N, Dhannoon, Huda M. Abass, "Printed Arabic Character Recognition Using Neural Network," Journal of Emerging Trend in Computing and Information Science, Vol 5, 2004, pp. 64-66.
- [6] Reza GharoieAhangar, Mohammad FarajpoorAhangar, "Handwritten Farsi Character Recognition Using Artificial Neural Network," International Journal of Computer Science and Information Security, Vol 4, 2009.
- [7] <http://www.omniglot.com/writing/igbo.htm>
- [8] <http://www.faculty.ucr.edu/~legneref/igbo/igbo2.htm>
- [9] Rafeal C. Gonzalez, Richard E. Woods, Digital Image Processing, second Edition, pp.596
- [10] Rashm, Mukesh Kumar, and Rohini Saxena, Algorithm And Tech-

nique On Various edge detection: a survey, Signal & Image Processing : An International Journal (SIPIJ) Vol.4, No.3, June 2013.

- [11] Robert Fisher, Simon Perkins, Ashley Walker, [http:// homepages.inf.ed.ac.uk/rbf/HIPR2/hipr_top.htm](http://homepages.inf.ed.ac.uk/rbf/HIPR2/hipr_top.htm),2004.
- [12]<http://www.itu.dk/courses/MBSB/E2005/Download/morphology.pdf>
- [13] Prof. PrimozPotocnik, ZigaZadnik "Character Recognition" pp 1-12
- [14] Daniel Graupe, "Principle of Artificial Neural Networks", 2nd Edition, pp. 1.
- [15] Ebenezer O. Olaniyi, Khasman Adnan,"Onset Diabetes Diagnosis Using Artificial Neural Network", In International Journal of Scientific and Engineering Research, vol 5, issue 10,pp.754-759,oct., 2014.
- [16] Simon Haykin, "Neural Networks,"A comprehensive Foundation, 2nd Edition, pp. 157,1999.

ABOUT THE AUTHOR

Ebenezer Obaloluwa, OLANIYI a current master student of Electrical/Electronic Engineering in Near East University, Lefkosa, via Mersin 10, Turkey. He is a member of centre of innovation for artificial intelligence, British university of Nicosia, Mersin 10, turkey. His research interests are: Neural Network, Image Processing, Signal Processing ,Computer Vision, Machine learning and Robotics.

E-mail:obalolu117@yahoo.com

O. OYEDOTUN is a member of Centre of Innovation for Artificial Intelligence, British University of Nicosia, Girne, via Mersin-10, Turkey and currently pursuing masters degree program in Electrical/Electronic Engineering in Near East University, Lefkosa,via Mersin-10, Turkey. Research interests include artificial neural networks, pattern recognition, machine learning, image processing, fuzzy systems and robotics.

PH.+905428892591.

E-mail: oyebade.oyedotun@yahoo.com

A. KHASHMAN received the B.Eng. degree in electronic and communication engineering from University of Birmingham, England, UK, in 1991, and the M.S and Ph.D. degrees in electronic engineering from University of Nottingham, England, UK, in 1992 and 1997. During 1998-2001 he was an Assistant Professor and the Chairman of Computer Engineering Department, Near East University, Lefkosa, Turkey. During 2001-2009 he was an Associate Professor and Chairman of Electrical and Electronic Engineering Department, and in 2009 he received the professorship. In 2001 he established the Intelligent Systems Research Group (ISRG) at the same university, and has been chairing ISRG until present. From 2007 until 2008 he was also the Vice-Dean of Engineering Faculty, later on from 2013 until 2014 the Dean of the faculty at the same university. Since 2014, Professor Khashman was appointed as the Founding Dean of Engineering.