International Journal of Engineering Research And Management (IJERM)
ISSN: 2349- 2058, Volume-07, Issue-05, May 2020

Simple Procedures to Create HSCS

Naseem Asad, Ziad Alqadi, Ismail Shayeb, Qazem Jaber, Jamil Al-Azzeh

Abstract— **Human speech signal is an important digital data type and it si widely used in many vital applications such as computer security systems. In this paper research we will focus on introducing a simple, accurate and highly efficient procedure to build human speech classification system. A method of speech signal features extraction method will be introduced, tested and adopted for the proposed procedure. A features data base will be built and used in ANN creation, training and running. The obtained ANN will be used as a classification tool.**

Index Terms— **Speech, features vector, ANN, MLBP, extraction time, training time, running time**

I. INTRODUCTION

Digital signals [1], [2] such human speeches [3], [4] and color images [5], [6], [7] are very important type of data because they are using in any vital applications such security systems and computer classification systems [4], [5]. Here in this paper we will introduce a procedure to build a human speech classification system (HSCS) [6], [7].

The speech signal, as it emerges from a speaker's mouth, nose and cheeks, is a one-dimensional function (air pressure) of time. Microphones convert the fluctuating air pressure into electrical signals, voltages or currents, in which form we usually deal with speech signals in speech processing, speech signal is emerges from a speaker's mouth, nose and cheeks, is a one-dimensional function (air pressure) of time [1], [8],[9]. Microphones convert the fluctuating air pressure into electrical signals, voltages or currents, in which form we usually deal with speech signals in speech processing [10], [11], [12]. Human speech is an analogue signal which can be converted to digital signal by applying sampling and quantization as shown in figure 1.

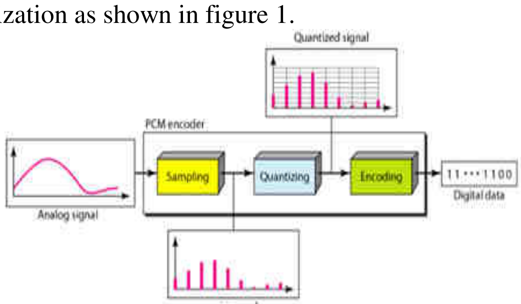

Figure 1: Converting speech analogue signal to digital

Manuscript received May 09, 2020

Naseem Asad, Princess Alia University College; Al-Balqa Applied University
Ziad Alqadi, Department of Computer Engineering. Albalqa Applied University, Faculty of engineering technology, Jordan, Amman
Ismail Shayeb, Princess Alia University College; Al-Balqa Applied University
Qazem Jaber, Department of Mechatronics Engineering, A1 Balqa Applied University
Jamil Al-Azzeh, Department of Computer Engineering. Albalqa Applied University, Faculty of engineering technology, Jordan, Amman

Speech signal is an important digital data type due to the vital applications requiring this kind of data, these applications such as security systems application [3], [4] require a high speed of implementation, but the speech signals usually have a big size, and thus will negatively affects the system efficiency and here we will seek a method to represent the speech by a small number of values to increase the process of speech manipulation. Speech signal file size depends on the recording time and the sampling rate [7], [8]. The sampling frequency or sampling rate, f_s, is the average number of samples obtained in one second (samples per second), thus $f_s = 1/T$. Table 1 shows some information about the speech signals which we will investigate in this paper [7], [8], [9].

Table 1: Used speech signal files

Speech #	Spoken words	Fs	Time(seconds)	Size(samples)	Size(bytes)
1	Please show me the cat image	44100	3.5331	155809	1246472
2	Please show me the dog image	44100	3.6465	160809	1286472
3	Please show me gorilla image	44100	3.3905	149520	1196160
4	Please show me the horse image	44100	3.5936	158477	1267816
5	Please show me the tiger image	44100	3.5235	155386	1243088
6	Please show me the lion image	44100	3.6419	160608	1284864
7	Please show me the Giraph image	44100	3.5251	155458	1243664
8	Please show me the donkey image	44100	3.8668	170528	1364224
9	Please show me the dolphin image	44100	3.9902	175970	1407760

10	Please show me the gazelle image	44100	3.5437	156279	1250232
Average			3.6255	159880	1279100

From table 1 we can see that the average number of samples is big, so the average file size is also big, and this will lead to extra time to identify the speech, so we can represent the speech file by a histogram [12], [13],[14] of 256 values and with size equal 2048 bytes for each speech file[5], [6], [9].
To reduce the classification time we have to use signal features instead of using the signal [13], the features vector for each speech signal will be unique, simple, fixed and will contain a small number of values [14], [15], [16].

II. PROPOSED PROCEDURE

To build HSCS we have to apply the following steps:

Step 1: Building a features database
Many methods are now used to create a speech signal features, some of these methods use linear prediction coding (LPC) [17], other methods use the clustering principles based on k_mean [18], [19] and fuzzy clustering [20], [21], other methods use finite impulse response filter (FIR) coefficients to create signal features [22], [23], other use wavelet packet tree (WPT) decomposition [24], [25]. These methods create unique features for each speech signal but they require high features extraction time.

In our research paper we will introduce a modified method based on local binary pattern (LBP) method [26], [27], [28], [29], and we will refer to it as modified LBP (MLBP).
MLBP features can be extracted performing the following tasks:

A. Get the speech signal.
B. For both types of mono and stereo signals, reshape the signal into one row.
C. Initialize the features vector to zeros (4 elements vector).
D. For each sample in the speech signal apply the step shown in table 2:

Table 2: MLBP calculation

Samples	….	S(i-2)	S(i-1)	S(i)	S(i-1)	S(i-2)	…….
Values	…	-0.5	1	0	-0.8	1	…
		<=	<=				
Binary		1	0				
Decimal	2 So add 1 to the vector with index=2						

So this step can be summarized in the following actions:

1. Select the speeches and save them in one folder.
2. For each speech signal create the features.
3. Add features as one column to features matrix (features data base).
4. When ending save the features database(FDB)

STEP 2: Creating and training ANN
Artificial neural network is a power full computational tool [30], [31], it contains a set of fully connected neurons [32],[33] arrange in one or more layers as shown in figure 2 [34], [35]:

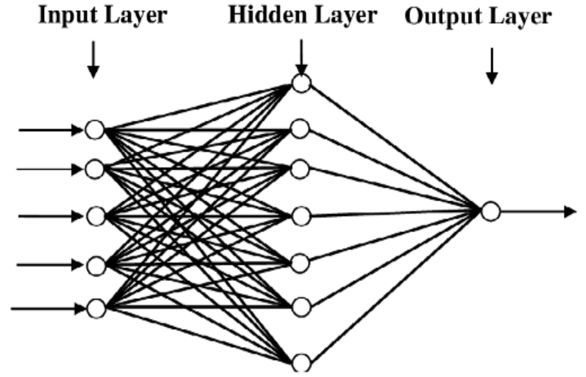

Figure 2: ANN with 3 layers

Each neuron acts as a computational element as shown in figure 3 [36], [37] and performs summation and output calculation according the activation function selected for this neuron, some of these functions are shown in figure 4.

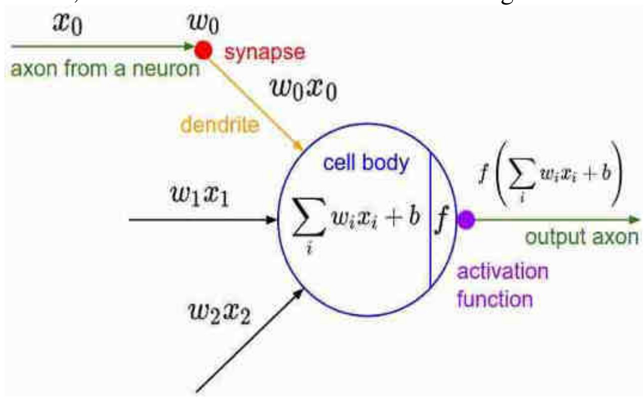

Figure 3: Neuron operations

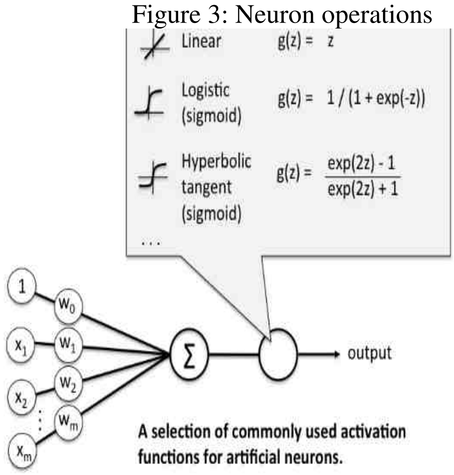

Figure 4: Some used activation functions

This step can be implemented applying the following actions:

1. Select FDB as ANN input
2. Define the targets, which are a numbers used as classifiers for the speech signals.
3. Creating ANN: here we have to select at least 2 layers: input layer with 4 neurons and output layer with I neuron, the activation function for the input layer must tansig or logsig, for the output layer it must be linear.
4. Initialize all the weights to zero.

5. Set the goal (error) to zero.
6. Select a number of training cycles (we use 500);
7. Train ANN using the inputs and the targets.
8. Check the outputs, if they are not acceptable adjust ANN architecture and train again.
9. If the outputs are acceptable save ANN to be used later.

Step 3: Running ANN as a classifier
In this final step we have use ANN as a classification tool by doing the following (for each run):

1. Get the speech signal features, by reading the signal and applying the same method used for features extraction.
2. Load ANN
3. Get the classifier by running ANN with the extracted features.
4. Do any action according to the classifier value,

Implementation and experimental results

1. Features extraction

The speech signals shown in table 1 were treated using MLBP method; table 3 shows the obtained results:

Table 3: Obtained speeches features

Speech #	Features				Extraction time(second)
1	43819	18251	1111	92624	0.0050
2	44592	18464	1024	96725	0.0070
3	40544	16775	948	91249	0.0040
4	44528	18834	1258	93853	0.0070
5	43203	17945	1055	93179	0.0080
6	46240	19029	1014	94321	0.0050
7	42981	18067	1151	93255	0.0060
8	48891	20251	1133	100249	0.0080
9	50043	21007	1287	103629	0.0110
10	43480	18072	1040	93683	0.0060
Average					0.0067

From table 3 we can see that the features for each signal are unique, the extraction process requires a significant small time (average=0.0067 seconds). For each speech signal the program was implemented several times and the features remain fixed without any changes.

At the end of this step we have to save the features which will be lock like figure 5:

```
Columns 1 through 8

 43819   44592   40544   44528   43203   46240   42981   48891
 18251   18464   16775   18834   17945   19029   18067   20251
  1111    1024     948    1258    1055    1014    1151    1133
 92624   96725   91249   93853   93179   94321   93255  100249

Columns 9 through 10

 50043   43480
 21007   18072
  1287    1040
103629   93683
```

Figure 5: Features DB

2. ANN creation and training

ANN was created using the procedures mentioned above, and then it was trained and gave acceptable outputs after 56 cycles as shown in figure 6:

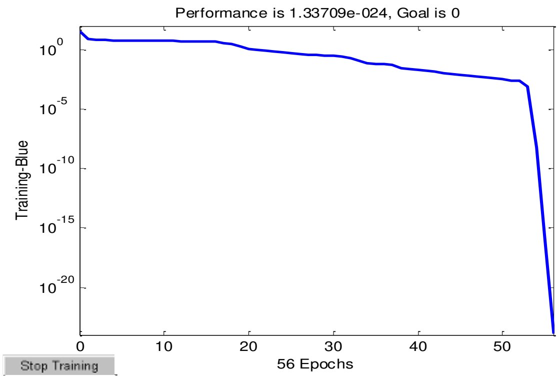

Figure 6: Training output

Here the creation and training cycle required 3.642000 seconds, but this cycle is to be performed once.

3. ANN running

The generated ANN was executed once for each speech signal, the recognition ratio was equal 100 %(each time of processing the calculated classifier was correct), the execution process requires a significant small time (in average 0.0922 seconds) as shown in table 4:

Table 4: Classification time

Speech number	Classification time(second)
1	0.120000
2	0.095000
3	0.090000
4	0.094000
5	0.088000
6	0.088000
7	0.079000
8	0.092000
9	0.08800
10	0.088000
Average	0.0922

Figures 7, 8, and 9 shows some classification outputs:

```
feat =

       43819
       18251
        1111
       92624

clas =

     1

Elapsed time is 0.120000 seconds.
type any key to continue.......
```

Figure 7: Getting the cat image

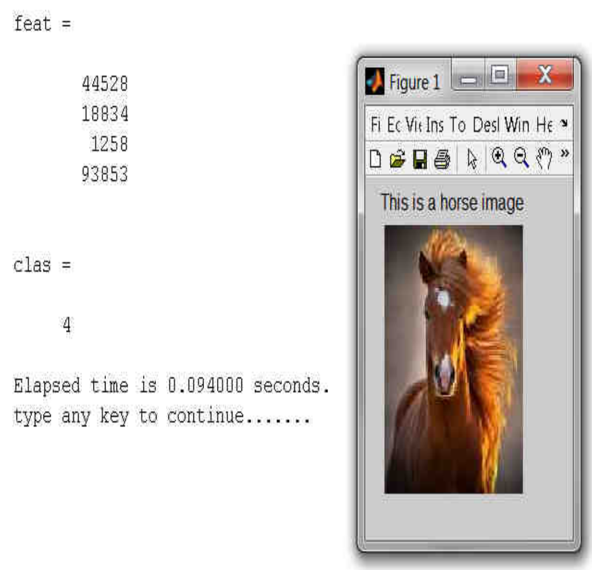

Figure 8: Getting the horse image

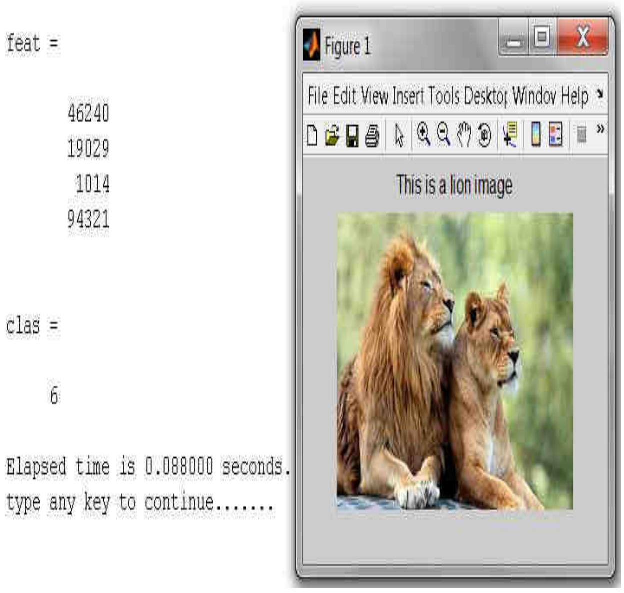

Figure 9: Getting the lion image

CONCLUSION

A simple, accurate and efficient procedure for building HSCS was introduced, the extracted features were simple, small, unique and fixed, the extraction time was significantly small, the proposed ANN gave 100 % accuracy and running ANN as speech classifier require a small amount of time. This procedure can adopted for any number of objects, thus we need some amendments for the featured data base and the targets.

REFERENCES

[1] Ziad Alqadi, Bilal Zahran, Jihad Nader, Estimation and Tuning of FIR Lowpass Digital Filter Parameters, International Journal of Advanced Research in Computer Science and Software Engineering, vol. 7, issue 2, pp. 18-23, 2017.

[2] Haitham Alasha'ary, Abdullah Al-Hasanat, Khaled Matrouk, Ziad Al-Qadi, Hasan Al-Shalabi, A Novel Digital Filter for Enhancing Dark Gray Images, European Journal of Scientific Research , pp. 99-106, 2014.

[3] Majed O Al-Dwairi, Ziad A Alqadi, Amjad A Abujazar, Rushdi Abu Zneit, Optimized true-color image processing, World Applied Sciences Journal, vol. 8, issue 10, pp. 1175-1182, 2010.

[4] Jamil Al Azzeh, Hussein Alhatamleh, Ziad A Alqadi, Mohammad Khalil Abuzalata, Creating a Color Map to be used to Convert a Gray Image to Color Image, International Journal of Computer Applications, vol. 153, issue 2, pp. 31-34, 2016.

[5] Musbah J. Aqel, Ziad A. Alqadi, Ibraheim M. El Emary, Analysis of Stream Cipher Security Algorithm, Journal of Information and Computing Science, vol. 2, issue 4, pp. 288-298, 2007.

[6] Jamil Al-Azzeh, Ziad Alqadi, Mohammed Abuzalata, Performance Analysis of Artificial Neural Networks used for Color Image Recognition and Retrieving, International Journal of Computer Science and Mobile Computing, vol. 8, issue 2, pp. 20 – 33, 2019.

[7] Akram A. Moustafa and Ziad A. Alqadi, Color Image Reconstruction Using A New R'G'I Model, Journal of Computer Science, vol. 5, issue 4, pp. 250-254, 2009.

[8] Dr. Ziad Alqadi, Akram Mustafa, Majed Alduari, Rushdi Abu Zneit, True color image enhancement using morphological operations, International review on computer and software, vol. 4, issue 5, pp. 557-562, 2009.

[9] Mohammed Ashraf Al Zudool, Saleh Khawatreh, Ziad A. Alqadi, Efficient Methods used to Extract Color Image Features, IJCSMC, vol. 6, issue 12, pp. 7-14, 2017.

[10] Ayman Al-Rawashdeh, Ziad Al-Qadi, Using wave equation to extract digital signal features, Engineering, Technology & Applied Science Research, vol. 8, issue 4, pp. 1356-1359, 2018.

[11] AlQaisi Aws, AlTarawneh Mokhled, A Alqadi Ziad, A Sharadqah Ahmad, Analysis of Color Image Features Extraction using Texture Methods, TELKOMNIKA, vol. 17, issue 3, 2018.

[12] Dr Ziad A AlQadi, Hussein M Elsayyed, Window Averaging Method to Create a Feature Victor for RGB Color Image, International Journal of Computer Science and Mobile Computing, vol. 6, issue 2, pp. 60-66, 2017.

[13] Dr. Ghazi. M. Qaryouti, Prof. Ziad A.A. Alqadi, Prof. Mohammed K. Abu Zalata, A Novel Method for Color Image Recognition, IJCSMC, Vol. 5, Issue. 11, pp.57 – 64, 2016.

[14] Jihad Nader Ismail Shayeb, Ziad Alqadi, Analysis of digital voice features extraction methods, International Journal of Educational Research and Development, vol 4, issue 1, pp. 49-55, 2019.

[15] Ahmad Sharadqh Naseem Asad, Ismail Shayeb, Qazem Jaber, Belal Ayyoub, Ziad Alqadi, Creating a Stable and Fixed Features Array for Digital Color Image, IJCSMC, vol. 8, issue 8, pp. 50-56, 2019.

[16] Majed O. Al-Dwairi, Amjad Y. Hendi, Mohamed S. Soliman, Ziad A.A. Alqadi, A new method for voice signal features creation, International Journal of Electrical and Computer Engineering (IJECE), vol. 9, issue 5, pp. 4092-4098, 2019.

[17] Dr. Amjad Hindi Dr. Majed Omar Dwairi Prof. Ziad Alqadi, PROCEDURES FOR SPEECH RECOGNITION USING LPC AND ANN, International Journal of Engineering Technology Research & Management, vol. 4, issue 2, pp. 48-55, 2020.

[18] Yousf Eltous Ziad A. AlQadi, Ghazi M. Qaryouti, Mohammad Abuzalata, ANALYSIS OF DIGITAL SIGNAL FEATURES EXTRACTION BASED ON KMEANS CLUSTERING, International Journal of Engineering Technology Research & Management, vol. 4, issue 1, pp. 66-75, 2020.

[19] Prof. Yousif Eltous, Dr. Ghazi M. Qaryouti, Prof. Mohammad Abuzalata, Prof. Ziad Alqadi, Evaluation of

Fuzzy and C_mean Clustering Methods used to Generate Voiceprint, IJCSMC, vol. 9, issue 1, pp. 75 -83, 2020.

[20] Ahmad Sharadqh Naseem Asad, Ismail Shayeb, Qazem Jaber, Belal Ayyoub, Ziad Alqadi, Creating a Stable and Fixed Features Array for Digital Color Image, IJCSMC, vol. 8, issue 8, pp. 50-56, 2019.

[21] Ahmad Sharadqh Jamil Al-Azzeh, Rashad Rasras , Ziad Alqadi , Belal Ayyoub, Adaptation of matlab K-means clustering function to create Color Image Features, International Journal of Research in Advanced Engineering and Technology, vol. 5, issue 2, pp. 10-18, 2019.

[22] Dr. Amjad Hindi Prof. Yousif Eltous Prof. Mohammad Abuzalata Prof. Ziad Alqadi Dr. Ghazi M. Qaryouti, USING FIR FILTER COEFFICIENTS TO CREATE COLOR IMAGE FEATURES, International Journal of Engineering Technology Research & Management, vol. 4, issue 2, pp. 6-14, 2020.

[23] Prof. Yousif Eltous Dr. Amjad Hindi, Prof. Ziad Alqadi, Dr. Ghazi M. Qaryouti, Prof. Mohammad Abuzalata, Using FIR Coefficients to Form a Voiceprint, International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering, vol. 8, issue 1, pp. 1-6, 2020.

[24] Dr. Ghazi M. Qaryouti, Prof. Mohammad Abuzalata, Prof. Yousf Eltous, Prof. Ziad Alqadi, Comparative Study of Voice Signal Features Extraction Methods, *IOSR* Journal of Computer Engineering (IOSR-JCE), vol. 22, issue 1, pp. 58-66, 2020.

[25] Amjad Y. Hindi, Majed O. Dwairi, Ziad A. AlQadi, Analysis of Digital Signals using Wavelet Packet Tree, IJCSMC, vol. 9, issue 2, pp. 96-103, 2020.

[26] Amjad Y. Hindi, Majed O. Dwairi, Ziad A. AlQadi, Creating Human Speech Identifier using WPT, International Journal of Computer Science and Mobile Computing, vol. 9, issue 2, pp. 117 – 123, 2020.

[27] Ismail Shayeb, Ziad Alqadi, Jihad Nader, Analysis of digital voice features extraction methods, International Journal of Educational Research and Development, vol. 1, issue 4, pp. 49-55, 2019.

[28] [27] Aws Al-Qaisi, Saleh A Khawatreh, Ahmad A Sharadqah, Ziad A Alqadi, Wave File Features Extraction Using Reduced LBP, International Journal of Electrical and Computer Engineering, vol. 8, issue 5, pp. 2780, 2018.

[29] Ziad Alqad, Prof. Yousf Eltous Dr. Ghazi M. Qaryouti, Prof. Mohammad Abuzalata, Analysis of Digital Signal Features Extraction Based on LBP Operator, International Journal of Advanced Research in Computer and Communication Engineering, vol. 9, issue 1, pp. 1-7, 2020.

[30] Ziad Alqadi, Aws Al-Qaisi, Adnan Manasreh, Ahmad Sharadqeh, Digital Color Image Classification Based on Modified Local Binary Pattern Using Neural Network, IRECAP, vol. 9, issue 6, pp. 403-408, 2019.

[31] [30] Ziad A AlQadi Amjad Y Hindi, O Dwairi Majed, PROCEDURES FOR SPEECH RECOGNITION USING LPC AND ANN, International Journal of Engineering Technology Research & Management, vol. 4, issue 2, pp. 48-55, 2020.

[32] Dr. Amjad Hindi, Dr. Majed Omar Dwairi, Prof. Ziad Alqadi, Analysis of Procedures used to build an Optimal Fingerprint Recognition System, International Journal of Computer Science and Mobile Computing, vol. 9, issue 2, pp. 21 – 37, 2020.

[33] Eng. Sameh S. Salma Prof. Ziad A. AlQadi, Eng. Ahmad S. AlOthma, Eng. Mahmoud O. Alleddaw ,Eng. Mohammad Ali Al-Hiar , Eng. Osama T. Ghaza, Investigation of ANN Used to Solve MLR Problem, IJCSMC, vol. 8, issue 5, pp. 38 – 50, 20, 2019.

[34] Eng. Sameh S. Salman Prof. Ziad A. AlQadi, Eng. Ahmad S. AlOthman, Eng. Mahmoud O. Alleddawi, Eng. Mohammad AlHiary, Eng. Osama T. Ghazal, Building Accurate and Efficient Color Image Recognizer, IJCSMC, vol. 8, issue 4, pp. 127 – 135, 2019.

[35] Jamil Al-Azzeh, Ziad Alqadi, Mohammed Abuzalata, Performance Analysis of Artificial Neural Networks used for Color Image Recognition and Retrieving, International Journal of Computer Science and Mobile Computing, vol. 8, issue 2, pp. 20-33, 2019.

[36] Belal Ayyoub, Ahmad Sharadqh, Ziad Alqadi, Jamil Al-azzeh, Simulink based RNN models to solve LPM, International Journal of Research in Advanced Engineering and Technology, vol. 5, issue 1, pp. 49-55, 2019.

[37] Abdullah Al-Hasanat, Haitham Alasha'ary, Khaled Matrouk, Ziad Al-Qadi, Hasan Al-Shalabi, Experimental Investigation of Training Algorithms used in Back propagation Artificial Neural Networks to Apply Curve Fitting, European Journal of Scientific Research, pp. 328-335, 2014.

[38] Ziad AA Alqadi, Investigation and analysis of mat lab models used to solve second order linear ODE, International Journal on Numerical and Analytical Methods in Engineering, vol. 2, issue 1, 2014.