International Journal on Recent and Innovation Trends in Computing and Communication
Volume: 3 Issue: 11

ISSN: 2321-8169
6220 - 6224

Application of Neural Networks in Sign Language Gesture Recognition

Ms. Rashmi S. Gaikwad
Electrical and Electronics Department
Sanjay Bhokare Group of Institutes, Miraj
India
e-mail- rashmigaikwad.rg@gmail.com

Mrs. Pooja P. Belagali
Department of Electronics & Telecommunication
J.J Magdum College of Engineering, Jaysingpur
India
e-mail- ppbelagali@rediffmail.com

Abstract— Neural Networks have a wide range of applications in the area of gesture recognition. This research gives a method to recognize English alphabets from A to Z in real-time using Neural Networks from a database of signs performed by a signer using a camera. A feature vector giving the dimensions of each sigh is calculated and processed using Neural Networks. The methodology is carried out using MATLAB 7.0 software.

Keywords- *Neural Network, Feature Extraction*

*****

I. INTRODUCTION

Many researches have been carried out to recognize hand gestures. The different methods used so far include Hidden Markov Models [6], Sequential search Algorithms etc. Some used colored gloves and biomechanical characteristics [1] to recognize hand gestures. Sign Language gesture recognition means recognition of signs of letters (in this case English Alphabets). In [9] the signs of Arabic Language have been carried out which follows the same methodology with slight differences.

Sign Language recognition will be helpful for the hearing impaired community to communicate with the normal people and normal people who do not know the sign language can use this method to convey their message. Thus, this research aims to reduce the communication gap between deaf people and normal people.

Fig- 1: Signs of English Alphabets

While designing this method many factors should be taken into consideration for example determining sign boundaries, occlusion, sensitivity, background condition and co-articulation that is mixing of preceding and subsequent images.

In this paper the first part gives the methodology of how the process is done. Next part gives the back propagation algorithm which is used in Neural Networks. The concluding parts gives the results and analysis of this method.

II. METHODOLOGY

Following Fig 1 shows the signs of all English letters from A to Z. A database is created which consists of total 26 signs which are performed by the signer.

The methodology of this research includes four stages (Fig 2) which can be summarized as data collection and image acquisition, image processing feature extraction and finally gesture recognition.

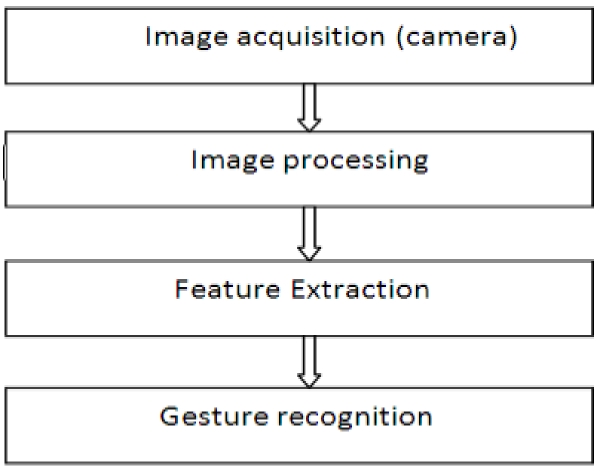

Fig -2: Design Steps

A. DATA COLLECTION AND IMAGE ACQUISITION

Stage one image acquisition is done using a digital camera and hand gestures are performed by participants. Next the feature extraction stage comes. Finally, stage four is concerned with using feed forward Neural Networks with back propagation for the gesture recognition purpose. Images were captured using a colored digital camera, each one of them was resized to 256 × 256 pixels, and image processing was performed using Matlab 7.10.

The image processing toolbox in Matlab represents color images as *RGB* (red, green, blue) values directly in an *RGB* image, which is then transformed to grey images. Some morphological operations are then applied to the grey images for removing deformities in them so that they can be applied for feature extraction.

B. FEATURE EXTRACTION

A parallelism exists between the feature extraction layer and the neural network layer. As a general practice, always the feature vector of the objects are derived all at a time and then sent as input to the first layer of network. The feature vector computation involves time and memory. In general, as the output classes are increased and the nonlinearity of differentiation in the classes increases, more feature vectors are necessary for the object recognition technique. The feature vector is a single row column matrix of N elements. These *n* elements are stored as shown in Fig. 3:

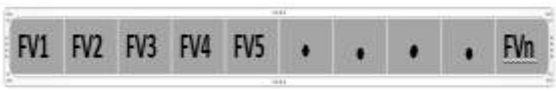

Fig- 3: Feature vector

The virtual LED display in matlab is considered to be 7x5. 35 features are extracted from the edge image of every alphabet sign and a database is created. The neural network will be trained with this database and when a test image is given as an input to the neural network, the neural network will compare the test image with the database and give the specific output. Figure 4 shows the output of the feature extraction stage which includes the input image, binary image, edge image and the alphabet image which is a 7x5 pixel array.

C. GESTURE RECOGNITION

Neural Network Architectures

Two main types of neural networks architectures can be used. Feed forward networks and recurrent networks. In this research, we have used four layered feed forward neural network with 16 neurons in the input layer, 50 neurons in the 1st hidden layer, 50 neurons in the 2nd hidden layer and 35 neurons in the output layer. More number of neurons results in increased performance.

Feed-Forward Neural Networks

Feed forward neural networks are the most popular and most widely used models in many applications. In this architecture each neuron in one layer has directed connections to the neurons of the subsequent layer; there are no links between neurons in the same layer neither with any of the previous layers [21]. The data flow in a strictly forward behavior. Since single-layer neural network has limited capabilities regarding pattern recognition; a multilayer feed forward neural network will be used.

Multilayer Feed forward Neural Networks

A multilayer feed forward network has a layered structure. The input layer where its neurons serve only as distribution points, one or more hidden layers of computation neurons, and the output layer [22].

Units in each layer receive their input from units from a layer directly below and send their output to units in a layer directly above the unit which means that the values only move from input to hidden to output layers; no values are fed back to earlier layers. Multilayer neural net-works have proven their ability to solve many difficult problems such as pattern recognition as well as the ability of to extract more meaningful features from the input patterns through the use of hidden layers [22].

The multilayer network architecture chosen for this research is a three layers neural network, *i.e.* a network that has one hidden layer. Since each gesture is represented by a vector containing thirty features; the input layer has been chosen to have 16 input units. There is no rule for determining the number of nodes that the hidden layer should have; many simulations lead us to decide on its number; 50 units. The output layer is 35 units since we have 35 feature vectors for every alphabet sign.

Each unit in the input layer has been fully connected to every other unit in the second layer—the hidden layer. Also, every unit in the hidden layer is connected to every other neuron in the output layer in a feed forward behavior.

The back-propagation Algorithm - a mathematical approach[23]

Fig 4 shows a Feed forward Neural Network. Units are connected to one another. Connections correspond to the edges of the underlying directed graph. There is a real number associated with each connection, which is called the weight of the connection. We denote by Wij the weight of the connection from unit ui to unit uj. It is then convenient to

ISSN: 2321-8169
6220 - 6224

represent the pattern of connectivity in the network by a weight matrix W whose elements are the weights Wij. Two types of connection are usually distinguished: excitatory and inhibitory. A positive weight represents an excitatory connection whereas a negative weight represents an inhibitory connection. The pattern of connectivity characterizes the architecture of the network.

A unit in the output layer determines its activity by following a two step procedure.

First, it computes the total weighted input xj, using the formula:

$$X_j = \sum_i y_i W_{ij} \tag{1}$$

where yi is the activity level of the jth unit in the previous layer and Wij is the weight of the connection between the ith and the jth unit.

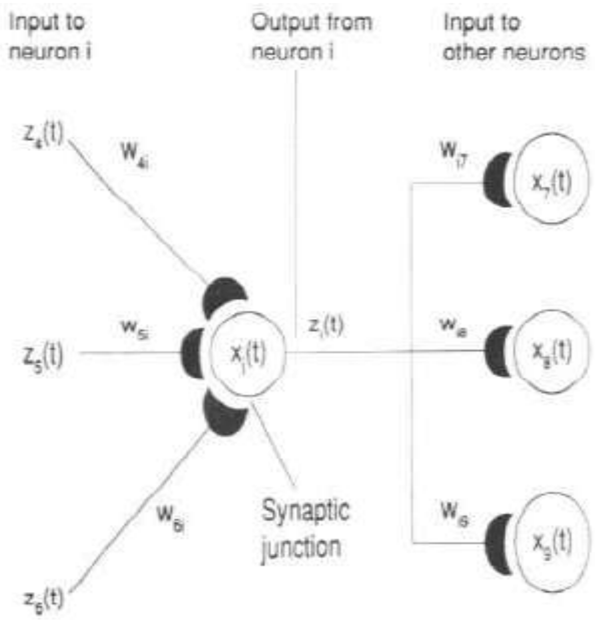

Fig- 4: Neural network

Next, the unit calculates the activity yj using some function of the total weighted input. Typically we use the sigmoid function:

$$y_j = \frac{1}{1 + e^{-x_j}} \tag{2}$$

Once the activities of all output units have been determined, the network computes the error E, which is defined by the expression:

$$E = \frac{1}{2}\sum_i (y_i - d_i)^2 \tag{3}$$

where y_i is the activity level of the ith unit in the top layer and d_i is the desired output of the jth unit.

The back-propagation algorithm consists of four steps

Compute how fast the error changes as the activity of an output unit is changed. This error derivative (EA) is the difference between the actual and the desired activity.

$$EA_j = \frac{\partial E}{\partial y_j} = y_j - d_j \tag{4}$$

Compute how fast the error changes as the total input received by an output unit is changed. This quantity (EI) is the answer from step 1 multiplied by the rate at which the output of a unit changes as its total input is changed.

$$EI_j = \frac{\partial E}{\partial x_j} = \frac{\partial E}{\partial y_j} \times \frac{\partial y_j}{\partial x_j} = EA_j y_j (1 - y_j) \tag{5}$$

Compute how fast the error changes as a weight on the connection into an output unit is changed. This quantity (EW) is the answer from step 2 multiplied by the activity level of the unit from which the connection emanates.

$$EW_{ij} = \frac{\partial E}{\partial w_{ij}} = \frac{\partial E}{\partial k_j} \times \frac{\partial k_j}{\partial W_{ij}} = EI_j y_i \tag{6}$$

Compute how fast the error changes as the activity of a unit in the previous layer is changed. This crucial step allows back propagation to be applied to multilayer networks. When the activity of a unit in the previous layer changes, it affects the activities of all the output units to which it is connected. So to compute the overall effect on the error, we add together all these separate effects on output units. But each effect is simple to calculate. It is the answer in step 2 multiplied by the weight on the connection to that output unit.

$$EA_i = \frac{\partial E}{\partial y_i} = \sum_j \frac{\partial E}{\partial k_j} \times \frac{\partial k_j}{\partial y_i} = \sum_j EI_j W_{ij} \tag{7}$$

By using steps 2 and 4, we can convert the EAs of one layer of units into EAs for the previous layer. This procedure can be repeated to get the EAs for as many previous layers as desired. Once we know the EA of a unit, we can use steps 2 and 3 to compute the EWs on its incoming connections.

III.ADVANTAGES OF NEURAL NETWORKS[24]

There are a variety of benefits that an analyst realizes from using neural networks in their work.

1) Pattern recognition is a powerful technique for harnessing the information in the data and generalizing about it. Neural networks learn to recognize the patterns which exist in the data set.

2) The system is developed through learning rather than programming. Programming is much more time consuming for the analyst and requires the analyst to specify the exact behavior of the model. Neural networks teach themselves the patterns in the data freeing the analyst for more interesting work.

3) Neural networks are flexible in a changing environment. Rule based systems or programmed systems are limited to the situation for which they were designed when conditions change, they are no longer valid. Although neural networks may take some time to learn a sudden drastic change, they are excellent at adapting to constantly changing information.

4) Neural networks can build informative models where more conventional approaches fail. Because neural networks can handle very complex interactions they can easily model data which is too difficult to model with traditional approaches such as inferential statistics or programming logic.

5) Performance of neural networks is at least as good as classical statistical modeling, and better on most problems. The neural networks build models that are more reflective of the structure of the data in significantly less time.

6) Neural networks now operate well with modest computer hardware. Although neural networks are computationally intensive, the routines have been optimized to the point that they can now run in reasonable time on personal computers. They do not require supercomputers as they did in the early days of neural network research.

IV. RESULTS AND ANALYSIS

The output of the algorithm can be connected to a 7x5 LED pixel array to indicate the letters. The advantages of this system are that the design is simple and the user does not require any special types of hand gloves to be worn. Figure 5 shows that the system was tested for various inputs and the corressponding signs were detected and the output is displayed using a text message as shown in Fig 5.

CONCLUSION

The signs for all the alphabets from A to Z are recognized using the combinational neural networks architecture. The images have been captured by a color camera and digitized into 256 × 256 pixel images. Image segmentation and the hand gesture recognition system are implemented using Matlab 7.10. The advantage of using the back propagation algorithm is high processing speed which produces results in real-time manner. The speed of processing is increased due to the use of neural network architecture. For future extensions processing of words and sentence gestures can be included.

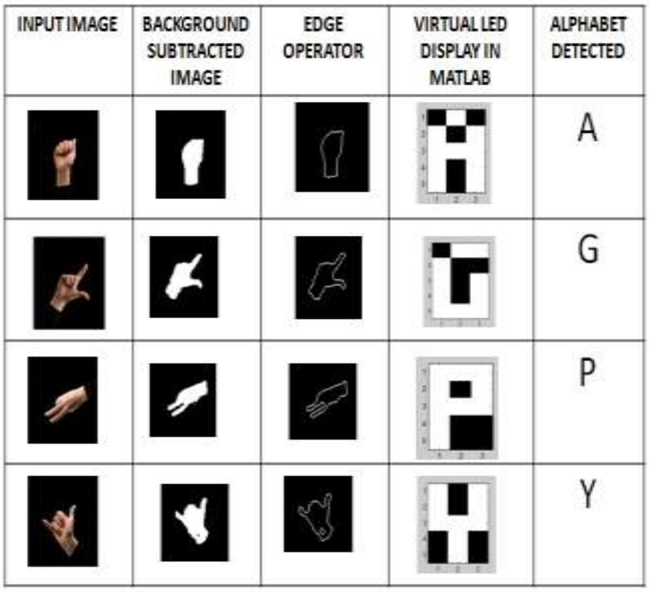

Fig- 5: Output

REFERENCES

[1] Farid Parvini, Dennis McLeod, Cyrus Shahabi, Bahareh Navai, Baharak Zali, Shahram Ghandeharizadeh "An Approach to Glove-Based Gesture Recognition" Computer Science Department University of Southern California Los Angeles, California 90089-0781 [fparvini,mcleod,cshahabi,navai,bzali,shahram]@usc.edu

[2] Paulraj M P, Sazali Yaacob, Hazry Desa, Hema C.R., " Extraction of Head & Hand Gesture Feature for Recognition of sign language", International Conference on Electronic Design, Penang, Malaysia, December 1-3, 2008.

[3] Paulraj M P, Sazali Yaacob, Mohd Shuhanaz bin Zanar Azalan, Rajkumar Palaniappan, "A Phoneme based sign language recognition system using skin color segmentation", Signal Processing and Its Applications (CSPA), pp: 1 – 5, 2010.

[4] Priyanka Mekala, Ying Gao , Jeffrey Fan , Asad Davari "Real-time Sign Language Recognition based on Neural Network Architecture" 978-1-4244-9592-4/11/$26.00 ©2011 IEEE

[5] Jong Bae Kim, Hye Sun Park,Min Ho Park,Massimo Piccardi,'Background subtraction techniques: a review',Systems, Man and Cybernetics,vol.4,2004,IEEE International Conference,pp:3099-3104.

[6] Honggang Hang , Ming C. Leuand Cemil oz , "American Sign Language Recognition Using Multi-dimensional Hidden Markov Models" Journal of Information Science and Engineering 22, 1109-1123 (2006)

[7] L. R. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," in *Proceedings of the IEEE*, Vol. 77, 1989, pp. 257-286.

[8] P. Mekala, S. Erdogan, Jeffrey Fan, "Automatic object recognition using combinational neural networks in surveillance networks", IEEE 3rd International Conference on Computer and Electrical Engineering (ICCEE'10), Chengdu, China, Vol. 8, pp. 387-391, November 16-18, 2010.

[9] Manar Maraqa, Dr. Raed Abu-Zaiter, "Recognition of Arabic sign language (ArSL) using recurrent neural networks", Applications of Digital Information and Web Technologies, pp: 478 – 481, 2008.

[10] K.Kawahigasi, Y.Shirai, J.Miura, N.Shimada "Automatic Synthesis of training Data for Sign Language Recognition Using HMM", PROC.ICCHP, pp.623-626, 2006.

[11] R. Liang and M. Ouhyoung (1998), "Real-time continuous gesture recognition system for sign language", Proc Third. IEEE International Conf: on Automatic Face and Gesture Recognition, pp. 558-567.

[12] P. Mekala, R. Salmeron, Jeffrey Fan, A. Davari, J. Tan, "Occlusion Detection Using Motion-Position Analysis", IEEE 42nd Southeastern Symposium on System Theory (SSST'10),Tyler, TX, pp. 197-201, March 7-9, 2010.

[13] Maebatake, M.; Suzuki, I.; Nishida, M.; Horiuchi, Y.; Kuroiwa, S., "Sign Language Recognition Based on Position and Movement Using Multi-Stream HMM", Universal Communication, 2008.ISUC '08. Second International Symposium, pp: 478 – 481, 2008.

[14] H.Brashear, K.-H.Park, S.Lee, V.Henderson, H.Hamilton, T.Starner, "American Sign Language Recognition in Game Development for Deaf Children" Proc.Assets, pp.79-86, 2006.

[15] Hong Pan and Liang-Zheng Xia, "Efficient Object Recognition Using Boundary Representation and Wavelet Neural Network", Neural Networks, IEEE Transactions, vol-19, pp.2132 - 2149.

[16] C. H. Lin and C. H. Wang, "Adaptive wavelet networks for power quality detection and discrimination in a power system," IEEE Trans. Power Delivery, vol. 21, no. 3, pp. 1106–1113, Jul. 2006.

[17] Y. C. Huang and C. H. Huang, "Evolving wavelet network for power transformer condition monitoring," IEEE Trans. Power Delivery, vol. 17, no. 2, pp. 412–416, Apr. 2002.

[18] K. H. Roberto, Galvao, and T. Yoneyama, "A competitive wavelet network for signal clustering," IEEE Trans. Syst. Man Cybern. B, Cybern., vol. 34, no. 2, pp. 1282–1288, Apr. 2004.

[19] Yang quan, "Chinese Sign Language Recognition Based on Video Sequence Appearance Modeling", Industrial Electronics and Applications (ICIEA), pp: 1537 – 1542, 2010.

[20] P. Vamplew, "Recognition of Sign Language Using Neural Networks," PhD Thesis, University of Tasmania, Newnham, 1996.

[21] J. Weissmann and R. Salomon, "Gesture Recognition for Virtual Reality Applications Using Data Gloves and Neural Networks," *Proceedings of IEEE International Joint Conference on Neural Networks*, Washington DC, 10-16 July 1999.

[22] S. Hykin, "Neural Networks: A Comprehensive Foundation," 2nd Edition, Prentice Hall, Upper Saddle River, 1998.

[23] James Momoh "Smart Grid: Fundamentals of Design and Analysis "Print ISBN: 9780470889398, Online ISBN: 9781118156117

[24] Amar Nath Singh, "Importance of Artificial neural network in modern Engineering" Journal of technological advances and Scientific Research, Sep.2015, Vol.1, Issue 3.