



International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijareeie.com

Vol. 6, Issue 3, March 2017

Multimodal Human Recognition

M.Valarmathi¹, Harshita Bhatt², Sutapa Banerjee³

Assistant Professor (O.G), Dept. of ECE, SRM University, Tamil Nadu, India ¹

Students, Dept. of ECE, SRM University, Tamil Nadu, India ^{2,3}.

ABSTRACT: Effective Security administration is one of the major issues of the present time. The efficient Security is best acknowledged when it is based on multiple factors. The customary method for security uses either human exertion or costly equipment. This work proposes a system in which human recognition is done by taking multiple factor (images and voice) as inputs using hardware which are economical and further processing is done by supervised algorithm which is an efficient means of enhance security. Along these lines, a successful framework for multimodal human recognition is made.

KEYWORDS: Beagleboneblack ,supervised algorithm ,Back propagation algorithm

I.INTRODUCTION

The efficient way to build security is best acknowledged when it is efficient, cost effective and based on multiple factors. Most of the security system in be implemented for highly secured environment as these system checks on the authorization of card not the person holding it. Thus these system often lead to security breaches. The other low budget systems includes human guard or system which checks on single attribute. So the it's the need of the hours to have a low budget system which can be easily implemented for secured environment This work proposes Multimodal human Recognition system which is fulfills all the above mentioned requirement an framework which recognize human by taking both voice and images. This system utilizes machine learning algorithm to recognize learns various input pattern and can be tested over new training pattern. This framework will be useful for organizations in securing their environment by using an autonomous and low budget effective tool.

II.LITREATURE REVIEW

Nowadays, a lot of researchers are working on the development of artificial neural networks. Neural networks. Some of them have implemented the human recognition using the generic algorithm. The disadvantage of using algorithm is that the update weights are random. Some of them has used fuzzy logic to implement human recognition but the disadvantage of using fuzzy logic is It is tedious to develop fuzzy rules and membership functions and fuzzy outputs can be interpreted in a number of ways making analysis difficult. In addition, it requires lot of data and expertise to develop a fuzzy system. It does not give generalizable results and the program has to be run for each individual node.High level efficiency is being implemented using back propagation and feed forward algorithms involving neural networks. Since it involved speech recognition, the epoch rate of multi-layered neural networks topology based on training set was computed. Several experiments were conducted with various weight.

III.IMPLEMENTATION

A. Input capturing frame work

The inputs capturing framework essentially comprises of a framework that can recognize human by taking multiple attribute to be specific:

1. Image
2. Voice

This input is taken by utilizing two sorts of hardware and processed by below mentioned processor :

1. Microphone
2. Logitech Camera



International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijareeie.com

Vol. 6, Issue 3, March 2017

3. Beagle Bone Black

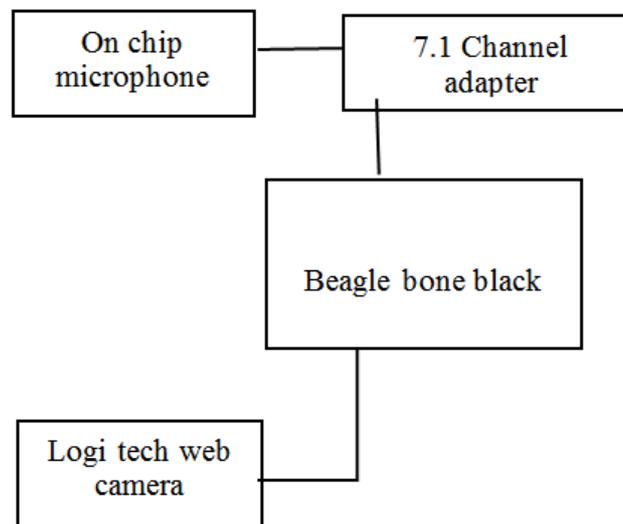
The Logitech web camera is used for capturing the images. The image capturing is triggered by the programming done on the beagle bone black. The python scrip has been used to trigger Logitech camera to capture images. Open cv is one of the python library used in the project to trigger camera to capture images. Since the project uses the machine learning supervised algorithm. The challenge to be resolved at this stage is environment adjustment. The camera should adjust the with the environment which is basically the light and brightness of the background. which has been resolved by coding in such a way that first three clicks are meant for camera environment adjustment. The images passed to the algorithm will be the one after third image capture. The web camera is interfaced with beagle bone black through USB cable.

B. Platform Used -Beagl bone black

Beagle bone black is a low cost development platform which has a capability of booting linux under 10 sec and starts working in less than 4 minutes using the usbcable. The processor used is AM335X ARM@ Cortex Rev C. The platform has the software compatibility with debian, Ubuntu, cloud9, android and much more. The connectivity to the platform can be achieved using USB host, Ethernet, hdmi and pin headers. The beagle bone black is a mini computer which has all the functionality like a normal pc

The logi tech camera is interfaced with the beagle bone black using USB cable. The on chip microphone is interfaced with beagle bone black using 7.1 channel adapter.

A supporting diagram of the project is given in Fig 1. It represents hardware setup.



C. Back propagation Algorithm

The Backpropagation consist of the two phases, feedforward path and back propagation path. Back propagation approximates the relationship between the input and the output by adjusting the weight matrix automatically. . Backpropagation network has two stages, training and testing. At the training phase, the network is provided with sample inputs and the corresponding classifications. The input in the proposed system is the image of the person the output is code that corresponds to the name of the respective person.

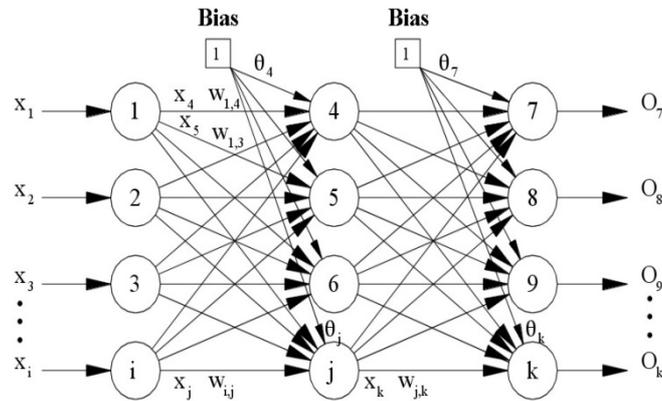
The following figure show representation of the Backpropagation algorithm which includes an input layer, hidden layer and an output layer

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijareeie.com

Vol. 6, Issue 3, March 2017



To the each of the node of the input layer .The hidden layer used for making the algorithm stable and for introducing modifications. The number of the nodes in the hidden layer could be any number.Each of the node in the input layer assign weight to the each of the node of the hidden layer.so the weighted sum of the input is given by following equation when training pattern is given to it.

$$Net_j = \sum w_{i,j} x_j + \theta_j$$

The theta term represents the bias node is which is a "pseudo input" to node in the hidden layer and the output layer, and is used to overcome the situations where the values of an input pattern are zero. If any input pattern has zero values, the system could not be trained without a this node.

$$out = x1 * (w1)^T$$

Similarly each of the node of the hidden layer assign weight to the each of the node of the output layer and the result will be given by the same formula- multiplication hidden layer input with the transpose of weights matrix at the output layer

$$net = [(x1 * (w1)^T) * (w2)^T]$$

The resulting output of the output layer is given to the activation function .The activation function used in the proposed project is sigmoid function which restrict output in the desired range in this particular case the output must not be negative and exceeds one thus the sigmoid function is used .

$$O_j = x_k = \frac{1}{1 + e^{-Net_j}}$$

The output of the activation fuction is the compared with the expected output in training phase .If the iteration is first time then the weight assigned would be randon.More error will be generated .So the aim of the backpropagation phase is to update the weight matrix in accordance to the error generated by the corresponding node .The error is given by

$$\Delta_k = t_k - O_k$$

The equation describes the overall error .the error corresponding to each of the contributing node is given by

$$\delta_k = \Delta_k O_k (1 - O_k)$$

The formulas used to modify the weight, $w_{j,k}$, between the output node, k, and the node, j is

$$\Delta w_{j,k} = Lr \delta_k x_j$$

Lr is the learning rate which is generally a small constant

In order to make the algorithm converge at the faster rate modification can be done in the weight update. Weight update during the nth iteration is given by including factor term (μ), which is multiplied with the (n-1)th iteration of the $\Delta w_{j,k}$. The introduction of the factor term is used to speed up the learning by pushing the weight changes to



International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijareeie.com

Vol. 6, Issue 3, March 2017

continue in the same direction with larger steps. In addition factor term prevents the learning process from striking in a local minima. By over stepping the small "hill".

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

IV. LIMITATIONS AND FUTURE SCOPE

Every proposed project has some constraints. By overcoming these constraints, we can further help in the development of the project. An idea of both the constraints as well as its future scope of the Multimodal Human Recognition is given as follows:

a. Limitations

Since the supervised algorithm is used the large number of training inputs are applied during the training phase

b. Future Scope

- Use of the exponential learning rate could be helpful in converging the network
- Modification can be made to back-propagation algorithm to increase the quality of a node by monotonic net incrementation. The training starts with a weight of few hidden nodes. Error containing trained node can be split periodically while learning the training input. The old weights are distributed by chance between the two new nodes. This is done until a maximum number of nodes in a hidden layer is reached. By training the inputs with the modified back-propagation algorithm a better minimum of the error is reached.

V. CONCLUSION

The system multimodal human recognition provides high level security by identifying human taking multiple factors as inputs that is both image and voice of the humans using back propagation algorithm

REFERENCES

- [1] Robert Barry, Artificial Neural Network Prediction of Wavelet Sub-bands for Audio Compression, Thesis for Bachelor of Science, Dept. of Electrical & Computer Engineering, 2000. <http://www.ilos.net/~rbarry/Thesis.pdf>
- [2] Raul Rojas, Neural Networks - A Systematic Introduction, Chapter 7, The Backpropagation Algorithm, 1996. <http://www.inf.fu-berlin.de/~rojas/neural/chap7.ps>
- [3] Ivan Galkin, U. MASS Lowell, Crash Introduction to Artificial Neural Networks
- [4] Based on a set of notes prepared by Dr. Patrick H. Corr, Techniques: The Multi-Layer Perceptron, <http://www.qub.ac.uk/mgt/intsys/backprop.html>
- [5] <http://www.cse.unsw.edu.au/~cs9417ml/MLP2/>