International Journal of Research Studies in Science, Engineering and Technology
Volume 1, Issue 5, August 2014, PP 65-71
ISSN 2349-4751 (Print) & ISSN 2349-476X (Online)

An Adaptive Learning for Recognizing Facial Expression from video using self Organizing Maps and Radial Basis Function Network

Abrham Debasu Mengistu, Prof. Vuda Sreenivasa Rao

M.Sc. (Computer Science), School of Computing and Electrical Engineering, BIT,
Bahir Dar University, Ethiopia.
Professor, School of Computing and Electrical Engineering, BIT,
Bahir Dar University, Ethiopia.

Abstract: R*ecognizing human facial expression in fully controlled environment is very simple task but in uncontrolled environments it becomes a challenging problem. i.e recognitions of facial expression of one's culture is simple but when we apply in uncontrolled environments it become difficult. Computer Vision has been researched for a long time, in the general case, there could be uncontrolled environments that may create error for the recognition of facial expression of humans. In this paper, we try to combine both supervised (RBF) and unsupervised learning network (SOM) method that can able to recognize human facial expression of different parts of the world.*

Keywords: *Artificial neural network, Facial recognition, SOM, RBF and Computer Vision*

1. Introduction

Machine learning involves adaptive mechanisms that enable computers to learn from the previous knowledge's. Learning capabilities can improve the performance of an intelligent system over time. The most popular approaches to machine learning are artificial neural networks. Artificial neural network or ANN is a machine learning techniques that is used to models human brain and consists of a number of artificial neurons. Neuron in artificial neural network consists of fewer connections than biological neurons each neuron in ANN receives a number of inputs and an activation function is applied to these inputs.

Neural networks look like the human brain in the following two ways:

- A neural network acquires knowledge through learning.
- A neural network's knowledge is stored within inter-neuron connection strengths known as synaptic weights

In artificial neural network a self organizing map or SOM for Unsupervised learning and radial basis function network or RBF for supervised learning are integrated in to one system for a computer vision that are capable of recognizing human activities from video. A learning machine need not be given all of the details of its environments if we provide all the details of the environment it takes time to execute all the instruction and its. Sensors are responsible for detecting the environments. In this paper we investigate neural network architecture for recognizing human facial activities that is taken from videos. A learning machine that is capable of recognizing human activities from videos is done by combining two learning paradigms which can be defined as follows:

Supervised learning , the goal is to construct an input and output mapping Y=F(x) that predicts the output y=(y1,y2…ym) for an input x=(x1,x2…xn) the mapping is found from example of the desired output {y(1),y(2)…} at the input data points (x(1),x(2)…} in order to minimize the expected output error.

Unsupervised learning only a set of input data {x(1),x(2)…}€ R is given and the goal is to construct a mapping so that the output {y(1),y(2)…} € R fully characterized the statistical properties of the input.

In general this requires a non linear function approximate with good generalization characteristics. We choose the radial basis function RBF network as the main framework for our study of recognizing human activities from video. Moreover, it is possible to train the RBF network in two stages with the basis function first being determined by unsupervised learning and then the second layer weights being determined by the supervised learning.

In computer vision for recognizing human facial activities system presented in this paper a self organizing map (SOM) is first used to cluster sensory information in to prototypes according to a topographic mapping then the RBF is used to implement the non linear mapping from sensory spaces to a motor action space. In this two layer learning structure, the SOM has two functions one is to compress the large amount of sensory information in order to reduce the computation time and the other is to divide the high dimensional sensor information space in to some small areas which are used to generate smoother approximations between the sensing information. The cluster centre obtained with the SOM are used to initialize the centre of the basis function in the RBF network and corresponding receptive fields are calculated by maximum like hood estimation the second layer weights in the RBF network are then obtained by the supervised learning. By using the well known least mean square (LMS) algorithm the output of the RBF Network are transitional.

2. Related Work

Artificial neural network is a popular approach in recent computer vision and robot system learning and used for simulating human brain. This requires learning of patterns from raw video and summarizing a video based on its content. Content based video summarization has been gaining renewed interest with corresponding advances in content based image retrieval (CBIR) [11]. Summarization and retrieval of consumer content such as sports videos is one of the most commercially viable applications of this technology [11].

Security and surveillance systems have traditionally relied on a network of video cameras monitored by a human operator who needs to be aware of the activity in the camera's field of view. With recent growth in the number of cameras and deployments, the efficiency and accuracy of human operators has been stretched. Hence, security agencies are seeking vision-based solutions to these tasks which can replace or assist a human operator. Automatic recognition of anomalies in a camera's field of view is one such problem that has attracted attention from vision researchers (c. f. [9], [10]). A related application involves searching for an activity of interest in a large database by learning patterns of activity from long videos [11], [12].

3. Statement of the Problem

In case of adaptive learning of recognizing human facial expression from video using RBF for controlled environment is very simple and straight forward because we feed all the training to the neural network then based on the training data set the network will recognize the expression. But when we apply for uncontrolled environments the capability of recognizing activities from video became minimal and cannot converge easily and also the convergence rate is very low.

4. Methodology

The methodology that we use for an adaptive learning of recognizing human facial expression activity from video is a combination of radial basis function (RBF) for recognizing in a fully controlled environment and SOM (Self organizing Map) for uncontrolled environments. In RBF all the training data set is given to the network for training and once the network is trained using RBF for fully controlled it is a very simple task to recognize the activity then the output of this controlled environments or RBF will be taken by SOM for uncontrolled environments this helps us to take a minimum iteration for choosing an activation value and also provides a higher rate of convergence.

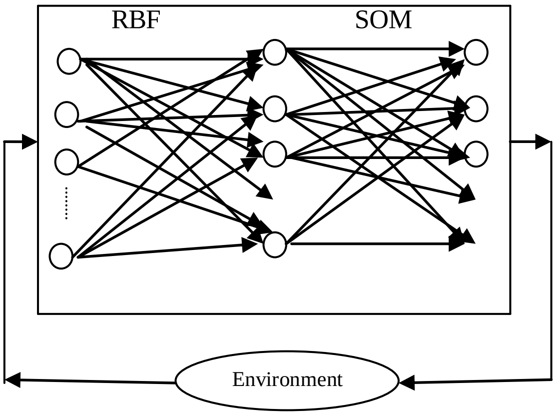

Figure 1. *SOM and RBF*

5. The Learning System

An adaptive learning system for an adaptive learning for recognizing human facial expressions from video through a learning computer vision system is composed of unsupervised learning and supervised learning; there are also input neurons at unsupervised learning. The architecture contains SOM for unsupervised learning and RBF for supervised are integrated in to one system for acquiring and recognizing human facial expressions of human which is sensed by a vision system. The major steps involved are the following:

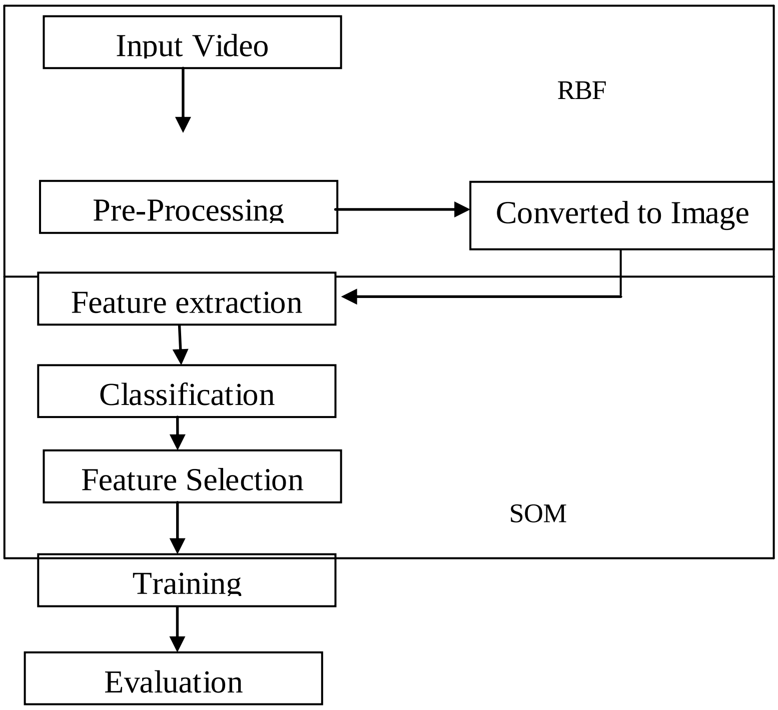

Figure 2. *An adaptive learning recognizing human facial expression*

5.1 RBF Neural Network Algorithm

1. Capture image from the Video
2. Convert to RGB
3. Choose the centers randomly from the captured image

Initialization: $t_k(0)$ randomk = 1, …, m_1

Sampling: draw x from input space

Similarity matching: find index of center closer to x

$$k(x) = \arg\min_k \|x(n) - t_k(n)\|$$

4. Compute the spread for the RBF function using the normalization method.

5. Find the weights using the pseudo-inverse method.

5.2 SOM Neural Network Algorithms

1. RBF output will be input to SOM
2. Initialize weights
3. for 0 to X number of training epochs
4. Select a sample from the input data sex find the wining neuron for a sample input image
5. Adjust the weights of nearby neurons
6. End for loop

6. EXPERIMENTS

Intensive experiments have been conducted to address the validity of the sensed videos. First, we have tested the vision program in several stages to check its accuracy. Initially, it is tested the achieved results from the vision program have a remarkable precision. Although the vision affected by the camera resolution, and lighting. As you see from figure 3: the training converges at 1200 epochs or iteration.

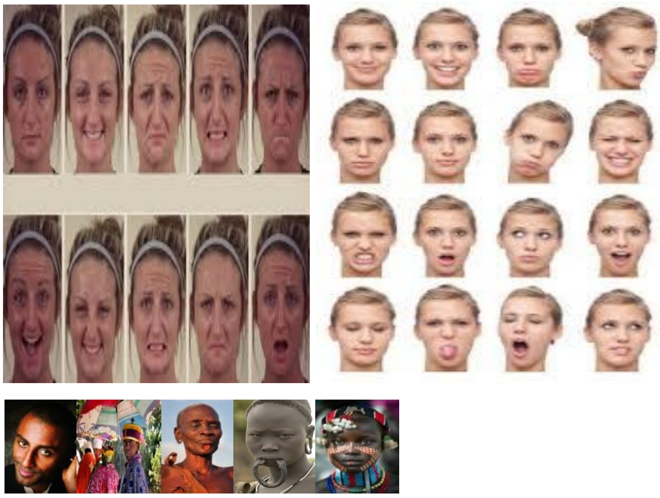

Figure 3. *Sensed Videos.*

6.1 Pre-Processing

Preprocessing is to prepare suitable images for analysis by performing feature enhancement and noise reduction. Preprocessing images commonly involves removing low-frequency background noise, normalizing the intensity of the individual particles images, removing reflections, and masking portions of images. After that, convert images from RGB to grey level where the features are based on grey level co occurrence matrix.

Figure 4. *Gray Scale Image*

6.2 Feature Extraction

The purpose of feature extraction is to reduce the original data set by measuring certain properties, or features, that distinguish one input pattern from another. All features extracted in this study are based on texture analysis using GLCM (gray level co-occurrence matrix). Suppose that the image of an algorithm is too large to be processed and it is suspected to be notoriously redundant then the input data will be transformed into a reduced representation set of features.

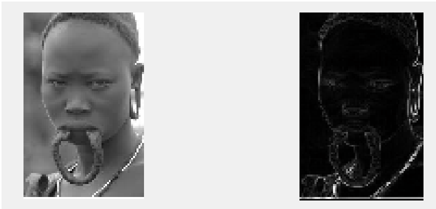

Figure 5. *Feature Extraction*

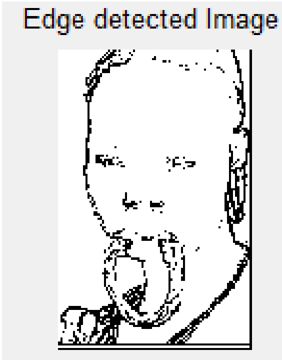

Figure 6. *Edge detected Image*

Figure 7. *Pattern Detection*

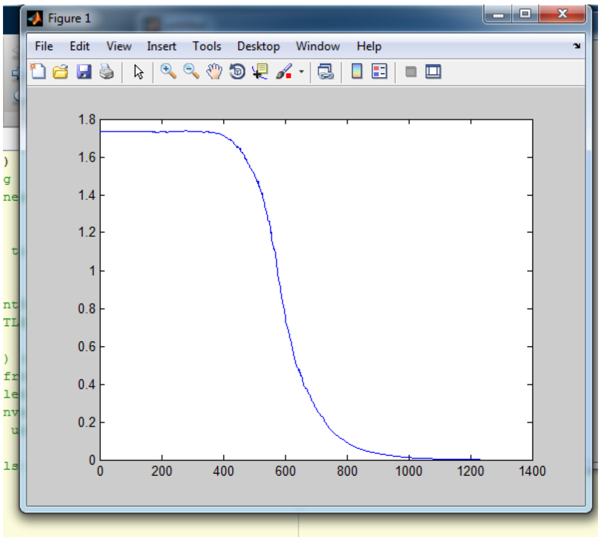

Figure 8. *Epochs*

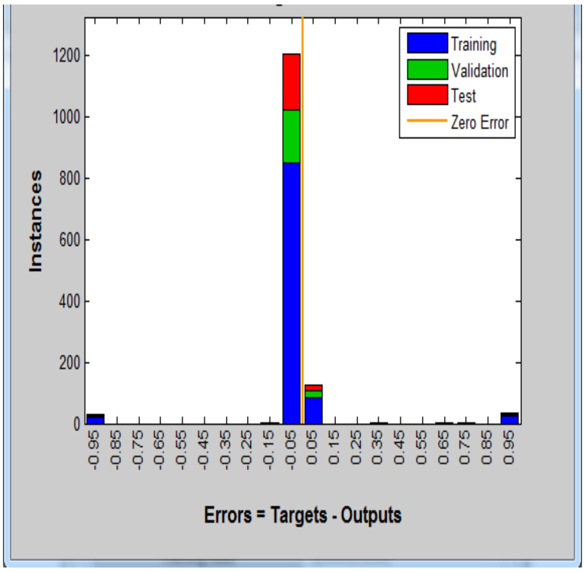

Figure 9. *Error on the target output*

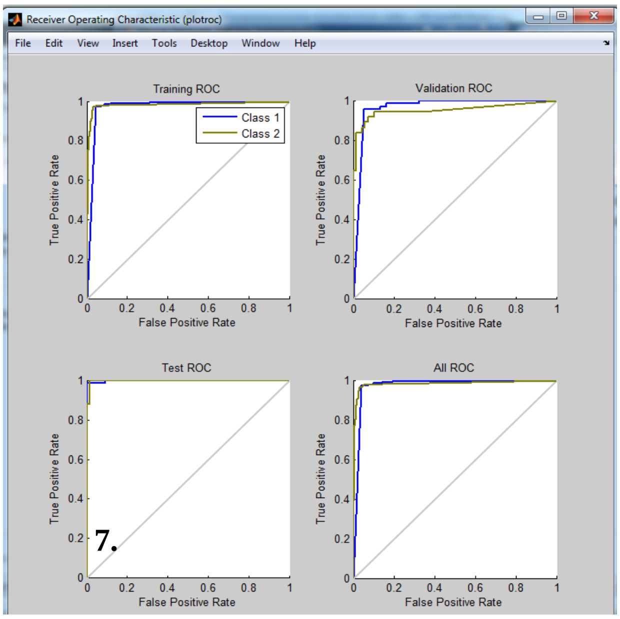

Figure 10. *Traget output*

8. Conclusion

Providing a machine the ability to see and understand as humans do has long fascinated scientists, engineers and even the common man. The researcher efforts in various scientific disciplines Computer Vision, AI, Neuroscience, Linguistics… etc has brought us closer to this goal than at any other point in history. However, several more technical and intellectual challenges need to be tackled before we get there. The advances made so far need to consolidated, in terms of their robustness to real world conditions and real time performance. This would then provide a firmer ground for further research.

References

[1] Haykin, S.Neural Networks: A comprehensive Foundation. 2nd Edition, Prentice Hall.

[2] M. Takizawa, Y. Makihara, N. Shimada, J. Miura and Y. Shirai, A Service Robot with Interactive Vision- Object Recognition Using Dialog with User

[3] http://www.gc.ssr.upm.es/inves/neural/ann1/anntutorial.html

[4] Paolo Caudiano, Eduardo Zalarna, and .Juan Lopez Coronado. An unsupervised neural network for real time low level control of a mobile robot: noise resistance, stability, and hardware implementation

[5] N. Vaswani, A. K. Roy-Chowdhury, and R. Chellappa, "“Shape Activity”: a continuous State HMM for moving/deforming shapes with application to abnormal activity detection,”

[6] Y. Rui, T. S. Huang, and S. F. Chang, “Image retrieval: current techniques, promising directions and open issues

[7] S. F. Chang, “The holy grail of content-based media analysis”

[8] H. Zhong, J. Shi, and M. Visontai, “Detecting unusual activity in video”

[9] C. Stauffer and W. E. L. Grimson, “Learning patterns of activity using real-time tracking”

[10] [10] W. Hu, D. Xie, T. Tan, and S. Maybank, “Learning activity patterns using fuzzy self-organizing neural network,”

[11] R. R. Elsley. Elsley. A learning architecture for control based on back-propagation neural networks.

AUTHORS’ BIOGRAPHY

Abrham Debasu Mengistu received his B.Sc. Degree in Computer Science from Bahir Dar University. Currently perusing M.Sc. in Computer Science, School of Computing and Electrical Engineering, IOT, Bahir Dar University, Ethiopia. He has published 01 research papers in international journal(IJSEAT). His main research interest is Image processing and Robotics. He is a life member of professional societies like MSDIWC.

Dr. Vuda Sreenivasa Rao received his M.Tech degree in computer science and engineering from Satyabhama University from 2007.He received PhD degree in computer science and engineering from Singhania University, Rajasthan, India from 2010. Currently working as Professor in School of Computing and Electrical Engineering, IOT, Bahir Dar University, Ethiopia. His main research interests are Data mining, Fuzzy logic, cloud computing and Network Security. He has got 14 years of teaching experience. He has published 41 research papers in various international journals and one Springer international conference paper. He has Editor-in-Chief in 3 international journals and 131 Editorial Board / Reviewer Board memberships in various international journals. He has Technical committee member in various international Conferences. He is a life member of various professional societies like IEEE, ACM, MAIRCC, MCSI, SMIACSIT, MIAENG, MCSTA, MAPSMS, MSDIWC, SMSCIEI, SNMUACEE and MISTE.