ADAPTIVE NEURAL NETWORK FOR SKETCH BASED IMAGE RETRIEVAL

Suchitra Agrawal
Department of CSE & IT
Madhav Institute of Technology and Science
Gwalior 474005, (India)

Rajeev Kumar Singh
Department of CSE & IT
Madhav Institute of Technology and Science
Gwalior 474005, (India)

Uday Pratap Singh
Department of Applied Mathematics
Madhav Institute of Technology and Science
Gwalior 474005, (India)

Abstract: In this paper, we present neural network approach for Sketch Based Image Retrieval (SBIR) using Histogram of Gradient (HOG) feature descriptor. This paper emphasis on back propagation Feed-forward Network (FFN) and Pattern Recognition Network (PRN) used for sketch based retrieval. Neural network is a popular tool used for pattern recognition and approximation of unknown nonlinear functions. We use features descriptor like Histogram of Gradient (HOG) for free hand and human face sketches and these features are used to training of network. Experimental results and analysis are based on CHUK and Flicker dataset, used for training and testing. Different similarity measure functions are discussed and used similarity between query by non-expert sketchers and database.

Keywords: FFN, SBIR, PRN, HOG, Sketch image.

1. INTRODUCTION

Two layer neural network has capability to retrieve similar images from big database to desired degree of accuracy with sufficient number of neurons in the hidden layer [1,2]. Neural network have been widely used for robust and adaptive controller to compensate the tracking error between query and retrieve images [3,4]. Neural network learning are successfully used as content based image retrieval system for many real world applications in which frequent changes are occurred due to environmental causes. In past few years different soft computing techniques have been used for image retrieval. Feed-forward network is one of the important network in which learning is supervised learning for retrieval of similar images. An excellent learning capabilities of fuzzy logic systems and neural network have been studied [5,6] for image retrieval systems. The key idea for sketch based image retrieval is based on extract edges of the sketch images using some image feature descriptor and then compare via some similarity/distance measure function between sketch feature and image features. The problem lies with the system is that the query image is a free-hand drawn sketch, the edge matching between the images is quite hard to provide the accurate results. A brief and concise list of research based on content or sketch based are described below.

A large number of scale feature vectors can be used for transforming image data [7] into scale-invariant co-ordinates. Corresponding features for each image/sketch pair are stored in a database and the query/image features are compared to find the matching. The general benchmark [8] for assessing the execution of any SBIR framework has been characterized. The framework that can powerfully create shadows [9] generated or extracted from huge database of images has been proposed. The descriptors are constructed in such a way that the sketch as well as color image have to undergo the same preprocessing level. A bag-of-regions [10] has been presented to construct a SBIR framework. This system encodes the eminent shapes on different levels of points of interest by considering as part of locales. To recognize similarities [11] between a hand drawn sketch and the images in a database is an important process. An image retrieval framework for free-hand sketches providing the interactive search [12] of image database depicting shape has been developed. A strategy to lessen the appearance gap [13] in SBIR was proposed for sketches and extracted edges of images are dealt with as combination of lines, establishing the framework for providing a better view on edges. Sketch-image pairs that hold minimum distance on the basis of feature vector are labeled as relevant otherwise irrelevant.

The remaining part of this paper been organized as follows: section 2 presents the sketch and image feature descriptors and FFN algorithm. We discuss experimental results and comparison with other existing network and methods in section 3, followed by conclusion and future work in section 4.

2. FEATURE DESCRIPTORS

In order to select the input/output vector for fuzzy counter propagation network for sketch based image retrieval system, we consider some important feature descriptors to retrieve similar image to the query image (free hand sketch) from database.

2.1 HOG Feature Descriptors

Histogram of Oriented Gradients (HOG): HOG is one of the simple yet powerful feature descriptor for facial images [15]. It forms a histogram on the basis of orientations of magnitude of gradient of each pixel.

$$\begin{bmatrix} g_x(x,y) \\ g_y(x,y) \end{bmatrix} = \begin{bmatrix} \frac{\partial I(x,y)}{\partial x} \\ \frac{\partial I(x,y)}{\partial y} \end{bmatrix}$$

(3)

$$\varphi = \tan^{-1}(\frac{g_y}{g_x})$$

(4)

$$\rho = (g_x^2 + g_y^2)^{0.5}$$

(5)

In this paper, we use HOG feature descriptor due to simplicity and effectiveness. Klare, Li and Jain [14], on

matching forensic sketches to mugshot photos shows that HOG has larger the distance between features of sketch and photos, therefore more suitable for retrieval.

2.2 Feed-forward Network

One hidden layer feed-forward neural network is shown in figure (1), which $X = \left\{ (x_i, d_i) \right\}_{i=1}^N$ where $x_i \in \Re^n$ and $d_i \in \Re^n$ d is the input and target vector of network.
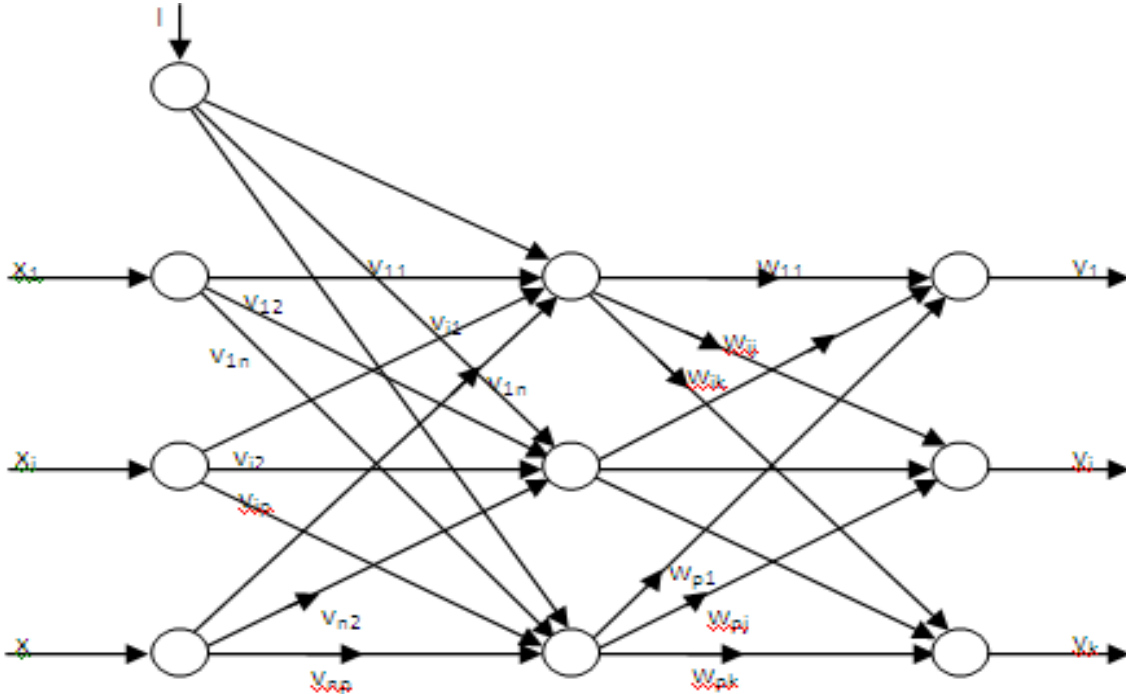

Figure (1): Architecture of feed-forward neural network

The neural network under consideration has general architecture shown in figure 1. Back Propagation Neural Network (BPNN) is multilayer (input, hidden and output) feed-forward network. Let us assume that the BPNN is homogeneous i.e. all the neurons have similar signal functions. For linear neuron in the input layer, $\phi(x) = x$ and for sigmoidal neurons in the hidden and output layers, $\phi(x) = \frac{1}{1+e^{-\lambda x}}$, where $\lambda \in (0,1]$. Let us assume that N training vector pairs, $X = \left\{ (x_i, d_i) \right\}_{i=1}^N$ where $x_i \in \Re^n$ and $d_i \in \Re^n$ d is the response vector of neural network when input x_i is presented as input to the network.

2.3 Error Function

Let vector pairs $X = \left\{ (x_i, d_i) \right\}_{i=1}^N$ are assumed to be samples representative of some nonlinear or unknown function $f \in \Re^n \to \Re^p$ which we have to approximate using neural network. The general idea is employ the gradient of the pattern error in order to reduce the global error over the entire training set. Change the weight of the network using error gradient of the pattern, so each weight change perturbs the existing neural network in order to

reduce the global pattern. If we denote initial neural network by $\aleph$, then sequential updation of this neural networks are denoted as, $\aleph^1, \aleph^2, \cdots, \aleph^i, \cdots$ where for a given input x_i neural network $\aleph^i$ generates an output signal $\phi(y_i)$ where y_i being the net input at the ith output neurons. The ith training pair (x_i, d_i) then define an instantaneous error is given by:

$$e_i = e_i - \phi(y_i)$$

(2)

where,

$$e_i = \left(e_1^i, e_2^i, \cdots, e_p^i \right)^T = \left(d_1^i - \phi\left(y_1^i \right), d_2^i - \phi\left(y_2^i \right), \cdots, d_p^i - \phi\left(y_p^i \right) \right)$$

(3)

The instantaneous summed of squared error E_i is the sum of square error of each individual output error e_k^i i.e.

$$E_i = \frac{1}{2} \sum_{k=1}^p \left(d_k^i - \phi\left(y_k^i \right) \right)^2 = \frac{1}{2} e_i^T e_i$$

(4) The mean square error *E*, is computed over the training set *X*, on neural network $\aleph$ as:

$$E = \frac{1}{N} \sum_{i=1}^{N} E_i$$

(5)

Learning Process of BPNN:
The basic steps of learning are outlined as follows:

1. select an input vector (sketch feature) x_i from training set *X*, and present it to the network.
2. compute activations and signals of input, hidden and output neurons.
3. compute the error over the output neurons by comparing the desired outputs (target feature vector).
4. use the error calculated in step 3, compute change in weights from input to hidden layer and compute change in weights from hidden to output layer, such that a global error measure gets reduced.
5. update change in weights of the network in accordance with the changes computed in step 4.
   from input to hidden layer weights:

$$v_{lj}^{i+1} = v_{lj}^{i} + \Delta v_{lj}^{i}$$

(6)

from input to hidden layer weights:

$$w_{jk}^{i+1} = w_{jk}^{i} + \Delta w_{jk}^{i}$$

(7)

6. repeat step 1-step 5 until the global error falls below a predefined limit.

Each sketch feature pattern presentation represents an *iteration* and a presentation of the entire training set is called *epoch*. In gradient descent learning weight changes are made in proportion to the negative direction of gradient and content of these error gradient are computed as $\frac{\partial E}{\partial v_{lj}^{i}}$ and $\frac{\partial E}{\partial w_{jk}^{i}}$, i.e. change in weights is computed from the MSE computed over the entire training set keeping neural network fixed. For input pattern $(x_1, x_2, \cdots, x_N)$ are sequentially presented to neural network $\aleph^1$, and calculate the global error using desired response and calculated neural network output, which is used to update the weights and generate $\aleph^2$ and so on.
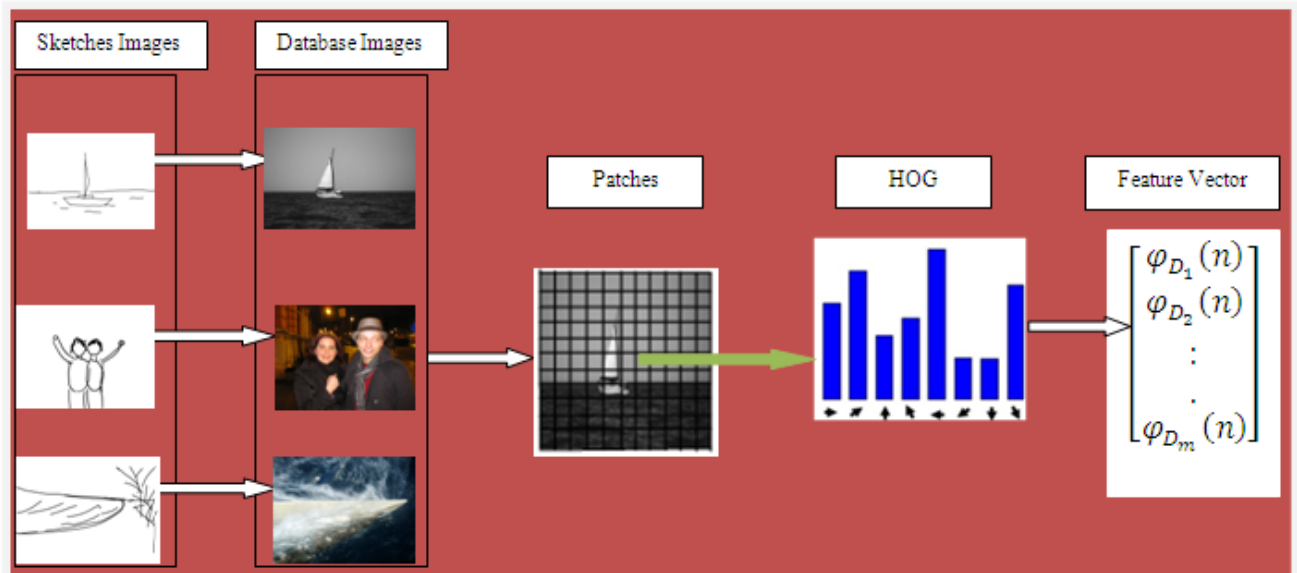

Figure 2: sketch feature and corresponding target feature vector using HOG

3. EXPERIMENTAL RESULTS

We present, experimental results using FFN approach on CHUK and Flickr dataset containing different categories of images. HOG feature vector is computed for sketches and database images. The vector thus generated serves as the input/output vector for the FFN and PRN. Results thus computed have been formed by performing multiple iterations to give a more general output. The parameters computed for achieving the efficiency and precision of the system has been manually calculated based on different factors such as: number of images to be retrieved, number of

relevant images present in the database for a certain sketch. Then, the average has been computed over the entire query sketches present.

$$Precision = \frac{Number\ of\ relevant\ images\ retrieved}{Total\ number\ of\ images\ retrieved}$$

On an average, system takes 1.511sec to retrieve the images containing similar object like sketch as shown in figure 2. Mean precision using different similarity measure function are shown in figure 3, and results obtained using FFN is shown in figure 4.
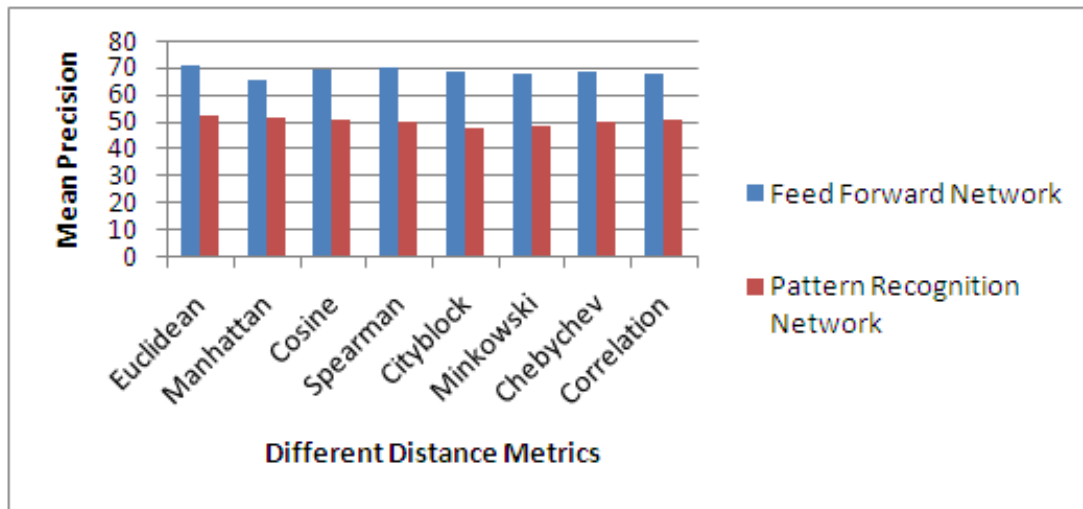
Figure 3: Comparison of FFN and PRN

4. CONCLUSION

In this paper, we present feed-forward neural network approach for sketch based image retrieval system. HOG is used for input and output feature vector of sketch/query image on CHUK and Flickr dataset comprises different categories of images. We have considered different distance measure to find the similarity between sketches and images. The experimental results demonstrate that feed-forward network for SBIR gives improved mean precision pattern recognition network. In future various image descriptors as well networks can be considered for sketch based image retrieval systems.
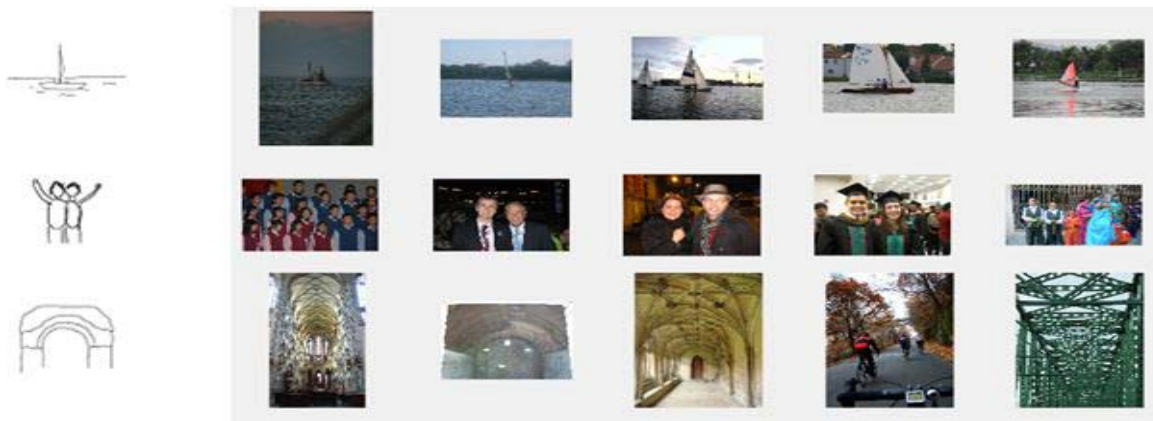

Figure 4: Comparison of (a) free hand sketch drawn by non expert sketchers and (b) similar photos recognised by FFN

REFERENCES

1. Uday Pratap Singh and Sanjeev Jain, "Optimization of neural network for nonlinear discrete time system using modified quaternion firefly algorithm case study of Indian currency exchange rate prediction", Soft Computing (Springer), DOI: 10.1007/s00500-017-2522-x, 2017.
2. K. Funahashi, "On the approximate realization of continuous mapping by neural networks," Neural Networks, vol. 2, pp. 183–192, 1989.
3. Uday Pratap Singh and Sanjeev Jain, "Modified Chaotic Bat Algorithm-Based Counter Propagation Neural Network for Uncertain Nonlinear Discrete Time System," International Journal of Computational Intelligence and Applications (World Scientific), vol. 15, no. 3, 2016, 1650016, DOI: 10.1142/S1469026816500164.
4. V. Sakhre, U. P. Singh, & S. Jain, "FCPN Approach for Uncertain Nonlinear Dynamical System with Unknown Disturbance," Int. J. Fuzzy Syst., vol. 19, no. 2, pp. 452-469, 2016. DOI:10.1007/ s40815- 016- 0145-5.
5. C. Ma, X. Yang, C. Zhang, X. Ruan and M. H. Yang, "Sketch retrieval via local dense stroke features", Image and Vision Computing, vol. 46, pp. 64-73, 2016.
6. F. Wang, L. Kang and Y. Li, "Sketch-based 3d shape retrieval using convolutional neural networks", IEEE Conference on Computer Vision and Pattern Recognition, pp. 1875-1883, 2015.
7. D. G. Lowe, "Distinctive image features from scale-invariant key points", International Journal of Computer Vision (Springer), vol. 60, no. 2, pp. 91–110, 2004.
8. M. Eitz, K. Hildebrand, T. Boubekeur and M. Alexa, "A descriptor for large scale image retrieval based on sketched feature lines", Proceedings of the 6th Eurographics Symposium on Sketch-Based Interfaces Modeling, pp. 29-36, 2009.
9. C. L. Zitnick, "Binary coherent edge descriptors", Proceedings of the 11th European Conference on Computer Vision: Part II, pp. 170-182, 2010.
10. R. Hu, T. Wang and J. Collomosse, "A bag-of-regions approach to sketch-based image retrieval", 18th IEEE International Conference on Image Processing, pp. 3661-3664, 2011.

11. K. Bozas and E. Izquierdo, "Large scale sketch based image retrieval using patch hashing", Advances in Visual Computing (Springer), pp. 210-219, 2012.

12. R. Hu and J Collomosse, "A Performance Evaluation of Gradient Field HOG Descriptor for Sketch Based Image Retrieval", Journal of Computer Vision and Image Understanding, Vol. 117, No. 7, pp. 790-806, 2013.

13. S. Wang, J. Zhang, T. X. Han and Z. Miao, "Sketch-Based Image Retrieval Through Hypothesis-Driven Object Boundary Selection With HLR Descriptor", IEEE Transactions on Multimedia, Vol. 17, No. 7, pp. 1045-1057, 2015.

14. B. Klare, Z. Li and A. K. Jain, "Matching forensic sketches to mug shot photos", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 33, No. 3, pp. 639–646, 2011.

15. O. D´eniz, G. Bueno, J. Salido and F. De la Torre, "Face recognition using histograms of oriented gradients", Pattern Recognition Letters, Vol. 32, No. 12, pp. 1598–1603, 2011.