Invariant Moments based War Scene Classification using ANN and SVM: A Comparative Study

Daniel Madan Raja S*
Senior Lecturer, Dept. of Information Technology,
Bannari Amman Institute of Technology, Sathyamangalam, Tamilnadu, India

Shanmugam A
Principal,
Bannari Amman Institute of Technology, Sathyamangalam, Tamilnadu, India

Abstract - In this paper we are trying to classify a war scene from the natural scene. For this purpose two set of image categories are taken viz., opencountry & war tank. By using Invariant Moments features are extracted from the images. The extracted features are trained and tested with (i) Artificial Neural Networks (ANN) using feed forward back propagation algorithm and (ii) Support Vector Machines (SVM) using radial basis kernel function with p=5. The comparative results are proving efficiency of Support Vector Machines towards war scene classification problems by using Invariant Moment feature extraction method. It can be concluded that the proposed work significantly and directly contributes to scene classification and its new applications. The complete work is experimented in Matlab 7.6.0 using real world dataset.
.

Keywords - Invariant Moments, Scene Classification, Artificial Neural Networks and Support Vector Machine

I. INTRODUCTION

Scene classification underlies many problems in visual perception such as object recognition and environment navigation. Scene and object classification are important research topics in robotics and computer vision. Computer Vision generally focuses on extracting what is where by merely looking at it. Many research problems have been studied and reported by the research community in the recent years. Scene classification refers to classifying the images into semantic categories (e.g. street, bedroom, mountain, or coast) [1], [2], [3]. Classification is one of the several primary categories of machine learning problems [4]. For the indoor - outdoor scene retrieval problem, the authors addressed how high-level scene properties can be inferred from classification of low-level image features [5]. Authors propose an automated method based on the boosting algorithm to estimate image orientations [6]. In [7], Bosch et al. present a scene description and segmentation system capable of recognizing natural objects (e.g., sky, trees, grass) under different outdoor conditions. In paper [8], the authors propose a new technique for the classification of indoor and outdoor images based on edge analysis. Analysis of texture [9] requires the identification of proper attributes or features that differentiate the textures of the image. Authors [10][11] analyze the efficiency of commonly used feature extraction methods such as haar features, invariant moments and co-occurrence matrix by using Artificial Neural Networks and Support Vector Machines classifiers for classifying natural scenes.

This paper presents the war scene classification using Invariant Moments feature extraction methods using Artificial Neural Networks with feed forward back propagation algorithm and Support Vector Machines with radial basis kernel function of p=5. The organization of the paper is as follows: Sections II describe Invariant Moments, Section III elaborates on Artificial Neural Networks, Section IV enunciates Support Vector Machines, Section V explains the proposed work, Sections VI & VII deal with implementation of ANN and SVM, Section VIII deals with discussion, and finally Section IX concludes with conclusion.

II. INVARIANT MOMENTS

Moment invariants are important shape descriptors in computer vision. The set of seven invariant moments ($\phi_1 - \phi_7$) was first proposed by Hu [12] for 2D images which was widely used contour-based shape descriptor. Two-dimensional moments of a digitally sampled M x M image that has gray function f(x,y) (x, y = 0,…,M-1) is given as,

$$m_{pq} = \sum_{x=0}^{x=M-1} \sum_{y=0}^{y=M-1} (x)^p \bullet (y)^q \, f(x,y)$$

where p, q = 0, 1, 2, 3… (1)

Seven invariant moments are:

$$\phi_1 = \eta_{20} + \eta_{02}$$
$$\phi_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2$$
$$\phi_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2$$
$$\phi_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2$$
$$\phi_5 = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]$$
$$\phi_6 = (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03})$$
$$\phi_7 = (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (3\eta_{12} - \eta_{30})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]$$

(2)

In particular, Hu [12] defines seven values, computed by normalizing central moments through order three, that are invariant to object scale, position and orientation. In terms of the normalized central moments, the seven moments are given (2).

III. ARTIFICIAL NEURAL NETWORKS

The first neurological network model was introduced by McCulloch and Pitts [13]. The Hebbian rule[14] represents neural learning procedures, which implies that the connection between two neurons is strengthened when both neurons are active at the same time. In [15], Werbos developed a learning procedure called backpropagation of error. Later on, the backpropagation of error learning procedure was separately developed and published by parallel distributed processing group [16], in which weights and biases are adjusted by error-derivative (delta) vectors backpropagated through the network. Backpropagation is commonly applied to feedforward multilayer networks. Sometimes this rule is called the generalized delta rule. Numerous ANN models are constructed; the differences in them might be the functions, the accepted values, the topology, the learning algorithms, etc.

In this work we use feed-forward artificial neural network using backpropagation algorithm. This is the most widely used neural network model, and its design consists of one input layer, at least one hidden layer, and one output layer as shown in “Fig. 1”. Each layer is made up of non-linear processing units called neurons, and the connections between neurons in successive layers carry associated weights. Connections are directed and allowed only in the forward direction, e.g. from input to hidden, or from hidden layer to a subsequent hidden or output layer. Back-propagation is a gradient-descent algorithm that minimizes the error between the output of the training input/output pairs and the actual network output.

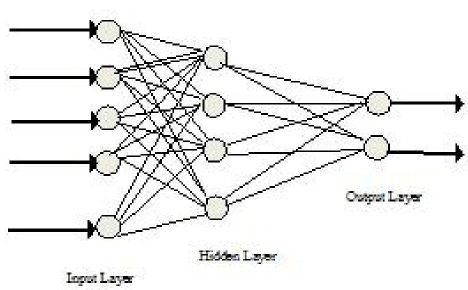


Figure 1. Simple Neural network Structure

Back propagation algorithm is applied for learning the samples, Tan-sigmoid and log-sigmoid functions are applied in hidden layer and output layer respectively, Gradient descent is used for adjusting the weights as training methodology.

IV. SUPPORT VECTOR MACHINES

Support Vector Machines are a new learning method for pattern recognition problem introduced by V.Vapnik et al [17][18]. An SVM classifies an input vector into one of two classes are based on the Structural Risk Minimization principle [17] from computational learning theory. The SVM learning algorithm directly seeks a separating hyperplane that is optimal by being a maximal margin classifier with respect to training data. Consider the problem of image classification where X is an input vector with 'n' dimensions. The SVM performs the following operation involving a vector $W = (w_1, ..., w_n)$ and scalar b:

$$f(X) = \mathrm{sgn}(W \bullet X + b) \tag{3}$$

Positive sign of f(X) may be taken as '*Opencountry*' images and negative value of f(X) may be regarded as 'War Tank' images. Consider a set of training data with l data points from two classes. Each data is denoted by (X_i, y_i), where i=1, 2,…,l, and $y_i \in \{+1, -1\}$. Note that y_i is a binary value representing the two classes. The graphical representation for a simple case of two-dimensional input (n=2) is illustrated in "Fig. 2".

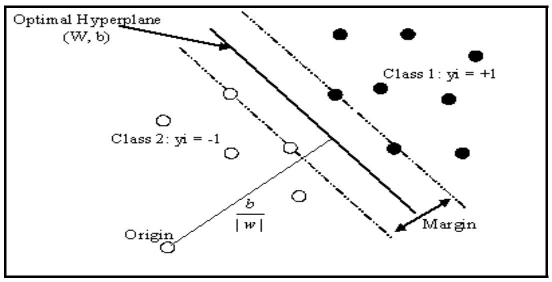


Figure 2. Optimal hyperplane for two class problem

Detailed discussions can be found in [18] [19] and [20]. Some commonly used kernel functions are:
Polynomial function:

$$K(X_i, X_j) = (X_i . X_j + 1)^d \tag{4}$$

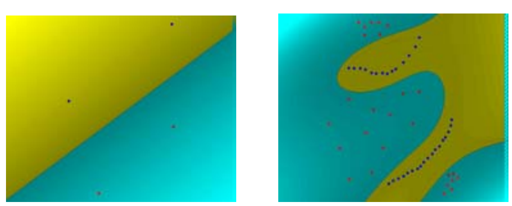
Radial basis function:

$$K(X_i, X_j) = e^{-\frac{\|x_i - x_j\|^2}{2\sigma^2}} \tag{5}$$

Sigmoid function :

$$K(X_i, X_j) = \frac{1}{1 + e^{[v(X_i . X_j) - \delta]}} \tag{6}$$

The hyperplane and support vectors used to separate the linearly separable data are shown in "Fig. 3" (a). And the hyperplane and support vectors used to separate the non-linearly separable data are shown in "Fig. 3" (b). Radial basis kernel function with p=5 used for this non-linear classification. Individual color represents particular class of data.



(a) (b)

Figure 3. (a) Linearly separable (b) Non-linearly separable

V. PROPOSED WORK

In classification, a classifier is trained to identify a type of example or differentiate between examples that fall in separate categories. In the case of computer vision, the examples are representations of photographic images and the task of the classifier is to indicate whether or not a specific object or phenomena of interest is present in the image. In order to successfully accomplish this, the classifier must have sufficient prior knowledge about the appearance of the image/scene. This paper is trying to recognize the scenes of two different categories called '*Opencountry*' and '*War tank scene*' *i.e. War tanks in opencountry*. The detailed work flow of the proposed system is shown in "Fig. 4".

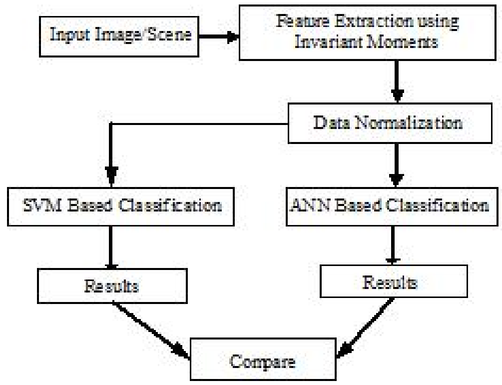


Figure 4. Detailed Description of Proposed Work

The sample images are taken from the Computational Visual Cognition Laboratory (opencountry) [21] and (War Tank scenes) is collected from the sources [22-31] with 200 samples each. Sample scenes are given in "Fig. 5" and "Fig. 6".

Figure 5. Sample images of 'Opencountry' category

Figure 6. Sample images of 'War Tank'' category

Invariant Moments are used for extracting the features from the images/scenes. The images are divided into four equal blocks and extracted seven values from each block. Thus, 4x7=28 features are used to represent an input image. Thus features F1 to F28 are considered as a feature set in invariant moments.

Normalization is then applied using Zero-mean normalization method in order to maintain the data within the specified range and also found suitable to improve the performance of the classifier.

VI. IMPLEMENTATION USING ANN

Using the above feature vector representations, neural classifier is trained and tested to recognize and classify the scenes. In Training phase, 200 samples are used including 100 samples from 'Opencountry' and 100 samples from 'War Tank Scenes'. In testing phase, 200 more samples are used including 100 samples from 'Opencountry' and 100 samples from 'War Tank Scenes'. The input images are converted into their gray scale images and resized to 256x256 pixels size. Zero-mean normalization method is applied to the extracted invariant moment features. Normalized features are given as input to Artificial Neural Networks to recognize the scene category. Backpropagation algorithm is used to train the neural classifier. The structure of the neural network is 28-8-2. "Fig.7" depicts the converging training graph of neural classifier.

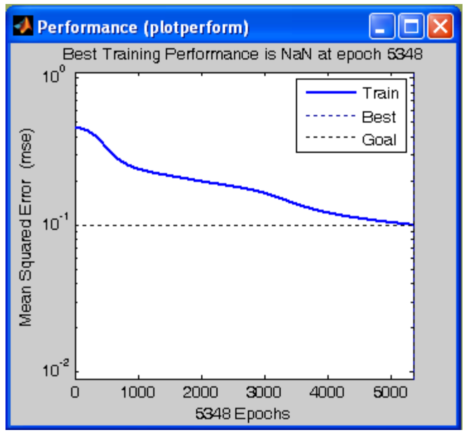


Figure 7. Converging training graph of neural classifier

VII. IMPLEMENTATION USING SVM

Using the above feature vector representations, Support Vector Machine is trained and tested to recognize and classify the scenes. In Training phase, 200 samples are used including 100 samples from 'Opencountry' and 100 samples from 'War Tank Scenes'. In testing phase, 200 more samples are used including 100 samples from 'Opencountry' and 100 samples from 'War Tank Scenes'. The input images are converted into their gray scale images and resized to 256x256 pixels size. Zero-mean normalization method is applied to the extracted invariant moment features. Normalized features are given as input to Support Vector Machine to recognize the scene category. Radial Basis Kernel Function with p=5 is used to train the classifier. The optimal hyperplane of trained radial basis kernel function that separates two different categories of data are shown in "Fig. 8".

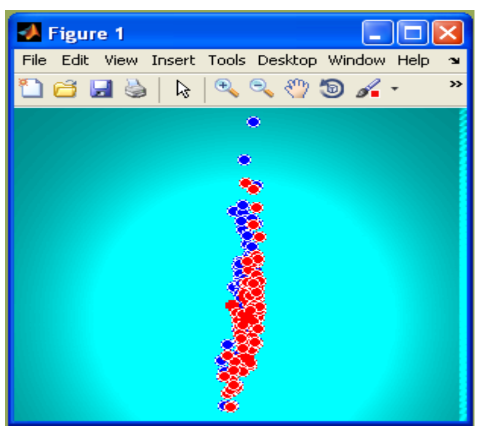


Figure 8. Optimal Hyperplane

VIII. **DISCUSSION**

This paper discusses invariant moment features based war scene classification using Artificial Neural Networks and Support Vector Machines. The sample images are taken from the Computational Visual Cognition Laboratory [21] and [22-31]. Features are extracted from the scene categories and the raw images are taken without any preprocessing steps to make the system robust to real scene environments. The pictorial representation which shows the comparative study of the performances of Artificial Neural Network and Support Vector Machines are shown in "Fig. 9" and "Fig. 10".

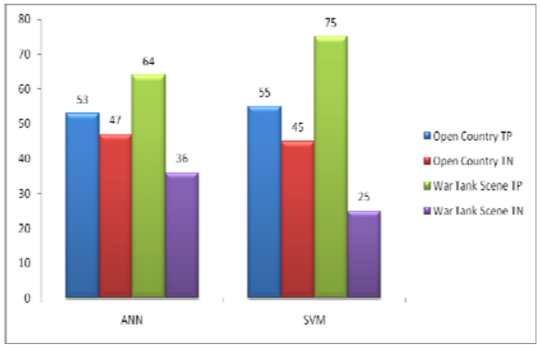


Figure 9. Performance of ANN and SVM

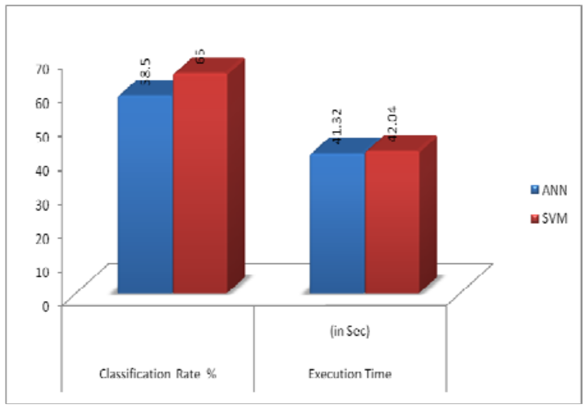


Figure 10. Classification Rate and Execution Time for ANN & SVM

The results show that invariant moment features based ANN is giving 58.5% classification rate in 41.32 seconds but invariant moment features based SVM is giving 65% classification rate in 45.34 seconds. The comparative results of ANN and SVM are given in table 1.

IX. CONCLUSION

This paper concentrates on the categorization of images as ‘War Tank’ scenes and ‘Opencountry’ scenes using invariant moments. The results are proving that invariant moments based SVM is giving higher classification rate i.e. 65% than ANN in war scene categorization problems. This work can be further extended to classify war scene categories using various feature extraction methodologies. The complete work is implemented using Matlab 7.6.0.

TABLE 1 Comparative Results of ANN and SVM

Feature Extraction Method	Classifier	Open Country		War Tank Scene		Classification Rate %	Execution Time (in Sec)
		TP	TN	TP	TN		
Invariant Moments	ANN	53	47	64	36	58.5	41.32
	SVM	55	45	75	25	65	42.04

TP=True Positive TN=True Negative

ACKNOWLEDGEMENT

The authors are extremely thankful to Computational Visual Cognition Laboratory and websites mentioned in the references [22-31].

REFERENCES

[1] Gokalp D., Aksoy S., "Scene Classification Using Bag-of-Regions Representations", Computer Vision and Pattern Recognition, CVPR, IEEE Conference on, (2007), pp.1-8.

[2] Vailaya A., A. Figueiredo, A. Jain, H. Zhang, "Image classification for content-based indexing", IEEE Transactions on Image Processing 10, (2001), pp.117–129

[3] Bosch A., A. Zisserman, and X. Mu˜noz, "Scene classification using a hybrid enerative/discriminative approach", IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 30, no. 4, (2008), pp. 712–727.

[4] Chella A., M. Frixione and S. Gaglio, "Understanding dynamic scenes", Artificial Intelligence 123 (2000) 89-132.

[5] Martin Szummer, Rosalind W. Picard, "Indoor-Outdoor Image Classification", Proceedings of the 1998 International Workshop on Content-Based Access of Image and Video Databases (CAIVD '98), January 03-03, (1998), p.42.

[6] Lei Zhang, Mingjing Li, Hong-Jiang Zhang, Boosting Image Orientation Detection with Indoor vs. Outdoor Classification, Proceedings of the Sixth IEEE Workshop on Applications of Computer Vision, December 03-04, (2002), p. 95.

[7] Bosch A., X. Munoz and J. Freixenet, "Segmentation and description of natural outdoor scenes", Image and Vision computing 25 (2007) pp. 727-740.

[8] Andrew Payne and Sameer Singh, "Indoor vs outdoor scene classification in digital photographs", Pattern Recognition 38 (2005) pp. 1533-1545.

[9] Arivazhagan S., L. Ganesan. Texture Segmentation Using Wavelet Transform. Pattern Recognition Letters, 24(16): December (2003), pp. 3197– 3203.

[10] Devendran V., Hemalatha Thiagarajan, A. K. Santra, Amitabh Wahi, "Feature Selection for Scene Categorization Using Support Vector Machines," Congress on Image and Signal Processing, Vol. 1, (2008), pp.588-592.

[11] Devendran V., Hemalatha Thiagarajan, Amitabh Wahi, " Texture based Scene Categorization using Artificial Neural Networks and Support Vector Machines: A Comparative Study" International Journal on Graphics, Vision and Image Processing, vol. 8,Issue IV, (2008), pp.45-52.

[12] Hu, M.K., "Visual pattern recognition by moments invariants", IRE Trans. Information Theory, 8,(1962), pp. 179-187.

[13] McCulloch, W. and W. Pitts, "A Logical Calculus of the Ideas Immanent in Nervous Activity", Bulletin of Mathematical Biophysics, Vol.5, (1943). pp. 115-133.

[14] Donald O. Hebb. The Organization of Behavior: A Neuropsychological Theory. Wiley, New York, 1949.

[15] Werbo P J., Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences. PhD thesis, Harvard University, 1974.

[16] Rumelhart D., G. Hinton, and R. Williams. Learning representations by back-propagating errors. Nature, 323:, October 1986, pp. 533–536

[17] Cortes C., and V. Vapnik, Support-vector networks. Machine Learning, 20:, November 1995, pp. 273-297.

[18] Vladimir N. Vapnik, The Nature of Statistical Learning Theory. Springer, New York, 1995.

[19] Vapnik, V.N., "The support vector method of function estimation. In: Generalization in Neural Network and Machine Learning. Springer-Verlag, New York, NY, 1998, pp. 239-268.

[20] Christopher J.C. Burges. "A Tutorial on Support Vector Machines for Pattern Recognition", Data Mining and Knowledge Discovery 2, (1998), pp. 121-167.

[21] http://cvcl.mit.edu/database.htm

[22] http://www.archives.gov/research/ww2/photos/

[23] http://www.militaryphotos.net/

[24] http://www.military.com/

[25] http://www.worldwar1.com/pharc.htm

[26] http://www.gwpda.org/

[27] http://www.historyofwar.org/

[28] http://en.wikipedia.org/wiki/Tanks_in_World_War_I

[29] http://en.wikipedia.org/wiki/Tanks_in_the_Cold_War

[30] http://en.wikipedia.org/wiki/Tanks_in_World_War_II

[31] http://en.wikipedia.org/wiki/Tank_classification

AUTHORS PROFILE

S.Daniel Madan Raja received B.E. Degree in Electronics and Communication Engineering from Government College of Engineering, Bargur in 2001 and M.E. Degree in Computer Science and Engineering from Hindustan College of Engineering, Chennai in 2003. He is working as a Senior Lecturer in the Dept. of Information Technology at Bannari Amman Institute of Technology, Sathyamangalam. He is doing a part time research in Anna University, Chennai. His current research focuses on Image Processing, Object Recognition, Pattern Recognition, Texture Analysis, Neural Networks and Support Vector Machines.

Dr.A.Shanmugam received B.E. Degree in PSG College of Technology in 1972, Coimbatore and M.E. Degree from College of Engineering, Guindy, Chennai in 1978 and Doctor of Philosophy in Electrical Engineering from Bharathiyar University, Coimbatore in 1994. From 1972–76, he worked as Testing Engineer in Testing and Development Centre, Chennai. He was working as a Lecturer Annamalai University in 1978. He was the Professor and Head of Electronics and Communication Engineering Department at PSG College of Technology, Coimbatore during 1999 to 2004. Authored a book titled “Computer Communication Networks” which is published by ISTE, New Delhi, 2000. He is currently the Principal, Bannari Amman Institute of Technology, Sathyamangalam.