

# ANN and SVM Based War Scene Classification Using Invariant Moments and GLCM Features: A Comparative Study

S. Daniel Madan Raja and A. Shanmugam

**Abstract**—Scene classification underlies many problems in visual perception such as object recognition and environment navigation. 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 and Gray Level Co-occurrence Matrix (GLCM), 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 GLCM feature extraction method. Although this study has been the first step of the research in the area of scene classification, the results have shown that the war scenes can be successfully classified from opencountry. 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.

**Index Terms**—Invariant moments, gray level Co-Occurrence matrix, scene classification, artificial neural networks, support vector machine.

## I. INTRODUCTION

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 and Gray Level Co-occurrence Matrix (GLCM) 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 & III describe Invariant Moments and Gray Level Co-occurrence Matrix (GLCM), Section IV elaborates on Artificial Neural Networks, Section V enunciates Support Vector Machines, Section VI explains the proposed work, Sections VII & VIII deal with implementation of ANN and SVM, Section IX deals with discussion, and finally Section X 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  $M \times M$  image that has gray function  $f(x,y)$  ( $x, y = 0, \dots, M-1$ ) is given as,

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

where  $p, q = 0, 1, 2, 3, \dots$

Seven invariant moments are:

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

Manuscript received June 10, 2012. Revised September 14, 2012

Daniel Madan Raja S is with the Department of Information Technology, Bannari Amman Institute of Technology, Tamil Nadu, India (e-mail: cross4u@rediff.com).

Shanmugam A is with the Department of Electronics and Communication Engineering, Bannari Amman Institute of Technology, Tamil Nadu, India (e-mail: dras\_bit@yahoo.com).

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. GRAY LEVEL CO-OCCURRENCE MATRIX

The procedure for extracting textural properties of image in the spatial domain was presented by Haralick *et al* [13]. The Gray Level Co-occurrence Matrix (GLCM) method considers the spatial relationship between pixels of different gray levels. The method calculates a GLCM by calculating how often a pixel with a certain intensity  $i$  occurs in relation with another pixel  $j$  at a certain distance  $d$  and orientation  $\theta$ . For instance, if the value of a pixel is 1 the method looks, for instance, the number of times this pixel has 2 in the right side. Each element  $(i, j)$  in the GLCM is the sum of the number of times that the pixel with value  $i$  occurred in the specified relationship to a pixel with value  $j$  in the raw image.

In this work, we are using eight texture descriptors i.e., energy; inertia; entropy; homogeneity; maxprob; contrast; inverse; correlation. Co-occurrence matrices are calculated for four directions:  $0^\circ$ ,  $45^\circ$ ,  $90^\circ$  and  $135^\circ$  degrees.

### IV. ARTIFICIAL NEURAL NETWORKS

The first neurological network model was introduced by McCulloch and Pitts [14]. The Hebbian rule [15] represents neural learning procedures, which implies that the connection between two neurons is strengthened when both neurons are active at the same time. In [16], 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 [17], 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.

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.

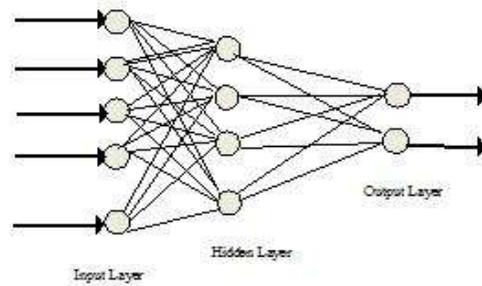


Fig. 1. Simple neural network Structure

### V. SUPPORT VECTOR MACHINES

Support Vector Machines are a new learning method for pattern recognition problem introduced by V. Vapnik *et al* [18] [19]. An SVM classifies an input vector into one of two classes based on the Structural Risk Minimization principle [18] 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  $W = (w_1, \dots, w_n)$  and scalar  $b$ :

$$f(X) = \text{sgn}(W \bullet X + b) \quad (3)$$

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  $i=1, 2, \dots, l$ , and  $y_i \in \{+1, -1\}$ . Note that  $y_i$  is a binary value representing the two classes. Detailed discussions can be found in [19] [20] and [21].

### VI. 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. 2.

The sample images are taken from the Computational Visual Cognition Laboratory (opencountry) [22] and (War Tank scenes) is collected from the sources [23-32] with 200 samples each. Sample scenes are given in Fig. 3 and Fig. 4.

Invariant Moments and Gray Level Co-occurrence Matrix are used for extracting the features from the images/scenes. In the case of invariant moments, images are divided into four equal blocks and extracted seven values from each block. Thus,  $4 \times 7 = 28$  features are used to represent an input image. Thus features F1 to F28 are considered as a feature set in invariant moments. In GLCM, eight features are extracted from each of the angular directions, 00, 450, 900 and 1350 degrees. So, total of 32 features are computed for each image. Thus features F1 to F32 are considered as a feature set in GLCM.

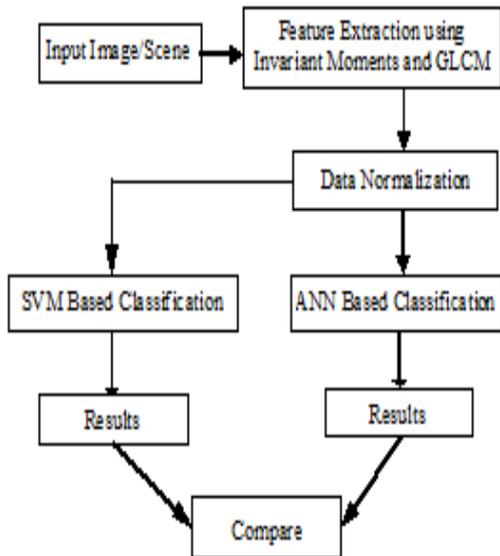


Fig. 2. Detailed description of proposed work

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.



Fig. 3. Sample images of 'Opencountry' category



Fig. 4. Sample images of 'War Tank' category

VII. 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 and GLCM 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 and 32-8-2 for invariant moment and GLCM features.

VIII. 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 and GLCM 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.

IX. DISCUSSION

This paper discusses invariant moment and GLCM features based war scene classification using Artificial Neural Networks and Support Vector Machines. The sample images are taken from the Computational Visual Cognition Laboratory [22] and [23-32]. 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. 5" and "Fig. 6".

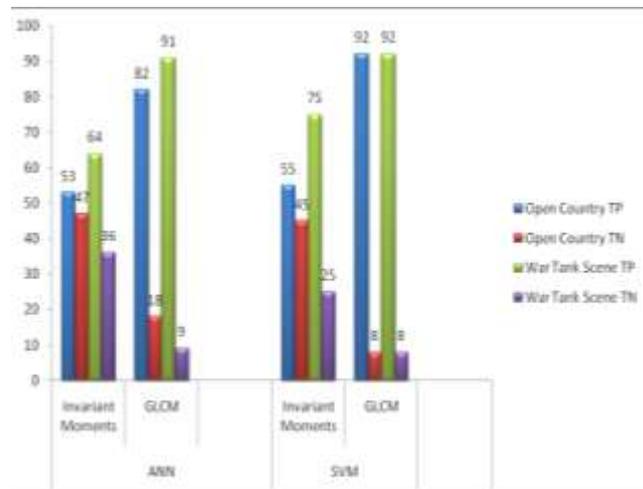


Fig. 5. Performance of ANN and SVM

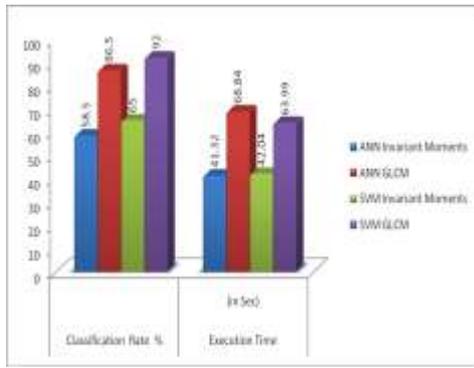


Fig. 6. Classification Rate and Execution Time for ANN & SVM

The results show that invariant moment features based ANN is giving 58.5% classification rate in 41.32 seconds, and GLCM features based ANN is giving 86.5% classification rate in 68.84 seconds. In the case of invariant moment features based SVM is giving 65% classification rate in 45.34 seconds and GLCM features based SVM is giving 92% classification rate in 63.99 seconds. The comparative results of invariant moment and GLCM are given in Table I.

TABLE I: COMPARATIVE RESULTS OF INVARIANT MOMENTS AND GLCM

Classifier	Feature Extraction Method	Open Country		War Tank Scene		Classification Rate %	Execution Time (in Sec)
		TP	TN	TP	TN		
ANN	Invariant Moments	53	47	64	36	58.5	41.32
	GLCM	82	18	91	9	86.5	68.84
SVM	Invariant Moments	55	45	75	25	65	42.04
	GLCM	92	8	92	8	92	63.99

TP=True Positive TN=True Negative

### X. CONCLUSION

This paper concentrates on the categorization of images as ‘War Tank’ scenes and ‘Opencountry’ scenes using invariant moments and GLCM features. The results are proving that GLCM features based ANN and SVM is giving higher classification rate i.e. 86.5% and 92% than invariant moments features based ANN and SVM 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.

### ACKNOWLEDGMENT

The authors are extremely thankful to Computational Visual Cognition Laboratory and websites mentioned in the references [23-32].

### REFERENCES

- [1] D. Gokalp and S. Aksoy, "Scene Classification Using Bag-of-Regions Representations," *Computer Vision and Pattern Recognition, CVPR*, IEEE Conference on, pp.1-8, 2007.
- [2] A. Vailaya, A. Figueiredo, A. Jain, and H. Zhang, "Image classification for content-based indexing," *IEEE Transactions on Image Processing*, vol. 10, pp.117-129, 2001.
- [3] A. Bosch, A. Zisserman, and X. Muñoz, "Scene classification using a hybrid generative/discriminative approach," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 30, no. 4, pp. 712-727, 2008.
- [4] A. Chella, M. Frixione and S. Gaglio, "Understanding dynamic scenes," *Artificial Intelligence*, vol. 123, pp. 89-132, 2002.
- [5] M. Szummer and 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), pp.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, pp. 95, 2002.
- [7] A. Bosch, X. Muñoz and J. Freixenet, "Segmentation and description of natural outdoor scenes," *Image and Vision computing* vol. 25, pp. 727-740, 2007.
- [8] A. Payne and S. Singh, "Indoor vs outdoor scene classification in digital photographs," *Pattern Recognition* vol. 38, pp. 1533-1545, 2005.
- [9] S. Arivazhagan and L. Ganesan, "Texture Segmentation Using Wavelet Transform," *Pattern Recognition Letters*, vol. 24, no. 16, December (2003), pp. 3197- 3203.
- [10] V. Devendran, H. Thiagarajan, A. K. Santra, and Amitabh Wahi, "Feature Selection for Scene Categorization Using Support Vector Machines," *Congress on Image and Signal Processing*, vol. 1, pp. 588-592, 2008.
- [11] V. Devendran, H. Thiagarajan and A. 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, no. 4, pp.45-52, 2008.
- [12] M. K. Hu, "Visual pattern recognition by moments invariants," *IRE Trans. Information Theory*, vol. 8, pp. 179-187, 1962.
- [13] R. M. Haralick, Shanmugam K, I. Dinstein, "Textural features for image classification," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 3, no. 6, pp. 610-621, 1973.
- [14] W. McCulloch and W. Pitts, "A Logical Calculus of the Ideas Immanent in Nervous Activity," *Bulletin of Mathematical Biophysics*, vol.5, pp. 115-133, 1943.
- [15] Donald O. Hebb., *The Organization of Behavior: A Neuropsychological Theory*. Wiley, New York, 1949.
- [16] P. J. Werbo, "Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences," PhD thesis, Harvard University, 1974.
- [17] D. Rumelhart, G. Hinton, and R. Williams. Learning representations by back-propagating errors. *Nature*, 323:, October 1986, pp. 533-536.
- [18] C. Cortes, and V. Vapnik, Support-vector networks. *Machine Learning*, 20:, November 1995, pp. 273-297.
- [19] Vladimir N. Vapnik, *The Nature of Statistical Learning Theory*. Springer, New York, 1995.

- [20] V. N. Vapnik, "The support vector method of function estimation," in *Generalization in Neural Network and Machine Learning*. Springer-Verlag, New York, NY, pp. 239-268, 1998.
- [21] Christopher J. C. Burges, "A Tutorial on Support Vector Machines for Pattern Recognition," *Data Mining and Knowledge Discovery* 2, pp. 121-167, 1998.
- [22] Computational Visual Cognition Laboratory, [Online]. Available: <http://cvcl.mit.edu/database.htm>
- [23] Pictures of World War II, [Online]. Available: <http://www.archives.gov/research/ww2/photos/>
- [24] Militaryphotos, [Online]. Available: <http://www.militaryphotos.net/>
- [25] Military, [Online]. Available: <http://www.military.com/>
- [26] Worldwar1, [Online]. Available: <http://www.worldwar1.com/pharc.htm>
- [27] The world war I documents achieves, [Online]. Available: <http://www.gwpda.org/>
- [28] History of war, [Online]. Available: <http://www.historyofwar.org/>
- [29] Tanks in World War I, [Online]. Available: [http://en.wikipedia.org/wiki/Tanks\\_in\\_World\\_War\\_I](http://en.wikipedia.org/wiki/Tanks_in_World_War_I)
- [30] Tanks in the Cold War, [Online]. Available: [http://en.wikipedia.org/wiki/Tanks\\_in\\_the\\_Cold\\_War](http://en.wikipedia.org/wiki/Tanks_in_the_Cold_War)
- [31] Tanks in World War II, [Online]. Available: [http://en.wikipedia.org/wiki/Tanks\\_in\\_World\\_War\\_II](http://en.wikipedia.org/wiki/Tanks_in_World_War_II)
- [32] Tanks classification, [Online]. Available: [http://en.wikipedia.org/wiki/Tank\\_classification](http://en.wikipedia.org/wiki/Tank_classification)



**Mr. S. Daniel Madan Raja** was born in India on July 1980. He received his BE degree in Electronics and Communication Engineering from Government College of Engineering, Bargur, Tamil Nadu, India in 2001 and his ME degree in Computer Science and Engineering from Hindustan College of Engineering, Chennai, Tamil Nadu, India in 2003. He is working as Assistant Professor (Sr. Grade) in the Department of Information Technology at Bannari Amman Institute of Technology, Sathyamangalam, Tamil Nadu, India. 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. He is currently the Principal, Bannari Amman Institute of Technology, Sathyamangalam.